```
           GenCore version 5.1.6
     Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    April 25, 2005, 00:09:27 ; Search time 1962.22 Seconds
                        (without alignments)
                        17301.659 Million cell updates/sec

Title:     US-10-613-728-1
Perfect score: 5735
Sequence:  1 ggatcctgcaaggtcacca.........ccataagagtttgagtcgac 5735

Scoring table: IDENTITY_NUC
           Gapop 10.0 , Gapext 1.0

Searched:  4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters:    8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    N_Geneseq_16Dec04:*
         1:   geneseqn1980s:*
         2:   geneseqn1990s:*
         3:   geneseqn2000s:*
         4:   geneseqn2001as:*
         5:   geneseqn2001bs:*
         6:   geneseqn2002as:*
         7:   geneseqn2002bs:*
         8:   geneseqn2003as:*
         9:   geneseqn2003bs:*
        10:   geneseqn2003cs:*
        11:   geneseqn2003ds:*
        12:   geneseqn2004as:*
        13:   geneseqn2004bs:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match	Length	DB	ID	Description
1	5735	100.0		5735		ADI28886	Adi28886 Inducible
2	5045	88.0		5443	12	AAD02408	Aad02408 Mouse alp
3	5045	88.0		5443	5	ABL49724	Abl49724 Mouse pro
4	1279.8	22.3		1679	2	AAT31006	Aat31006 Mouse car
5	1279.8	22.3		1679	6	ABV74357	Abv74357 Mouse car
6	1279.8	22.3		1679	12	ADL14241	Adl14241 Mouse car
7	407.6	7.1		5190	13	ADS97705	Ads97705 Rabbit al
8	311.8	5.4		520	2	AAT11359	Aat11359 Bidirecti
9	311.8	5.4		520	2	AAT45722	Aat45722 Bi-direct
10	311.8	5.4		520	2	AAV60079	Aav60079 Bidirecti
11	311.8	5.4		520	2	AAX60045	Aax60045 Bidirecti
12	311.8	5.4		520	3	AAZ56126	Aaz56126 Bidirecti
13	311.8	5.4		520	4	AAH47631	Aah47631 Bidirecti
14	311.8	5.4		520	5	AAH25572	Aah25572 Nucleotid
15	311.8	5.4		520	8	ACA94734	Aca94734 Tet coord
16	311.8	5.4		520	8	ABX16487	Abx16487 Bi-direct
c 17	311	5.4		4438	10	AAD63234	Aad63234 pTet07Sag
c 18	311	5.4		4479	10	AAD63233	Aad63233 pTet07Sag
c 19	311	5.4		4556	10	AAD63232	Aad63232 pTet07Sag
c 20	311	5.4		4556	11	ADO59258	Ado59258 pTet07-Sa
c 21	311	5.4	311	6346	10	AAD63231	Aad63231 pTet07Sag
c 22	311	5.4	311	6346	11	ADO59257	Ados9257 pTet07-Sa
c 23	311	5.4	311	6423	10	AAD63230	Aad63230 pTet07Sag
c 24	311	5.4	311	6424	11	ADO59253	Ados9253 pTet07-Sa
c 25	311	5.4	311	8287	10	AAD63235	Aad63235 pTet07Sag
c 26	311	5.4	311	8364	10	AAD63236	Aad63236 pTet07Sag
c 27	310.4	5.4		450	2	AAT06870	Aat06870 PhCMV*-2
c 28	310.4	5.4		450	2	AAQ76267	Aaq76267 PhCMV*-2
c 29	310.4	5.4		450	2	AAT11356	Aat11356 Minimal C
c 30	310.4	5.4		450	2	AAV60081	Aav60081 Cytomegal
c 31	310.4	5.4		450	2	AAX27903	Aax27903 PhCMV*-2
c 32	310.4	5.4		450	2	AAX81722	Aax81722 A human c
c 33	310.4	5.4		450	2	AAX60047	Aax60047 Minimal H
c 34	310.4	5.4		450	2	AAX01367	Aax01367 Nucleotid
c 35	310.4	5.4		450	3	AAZ56128	Aaz56128 Tetracycl
c 36	310.4	5.4		450	4	AAD09835	Aad09835 Human Cyt
c 37	310.4	5.4		450	4	AAH47633	Aah47633 Nucleotid
c 38	310.4	5.4		450	5	AAH25574	Aah25574 Nucleotid
c 39	310.4	5.4		450	5	AAL49541	Aal49541 Tet opera
c 40	310.4	5.4		450	8	ACA94736	Aca94736 Minimal C
c 41	310.4	5.4		450	8	ABX15339	Abx15339 Human Cyt
c 42	310.4	5.4		450	8	ABX16489	Abx16489 CMV promo
c 43	310.4	5.4		450	13	ADR88197	Adr88197 PhCMV*-2,
c 44	309.4	5.4		447	4	AAD09962	Aad09962 Tet respo
c 45	309.4	5.4		450	2	AAT06869	Aat06869 PhCMV*-1

ALIGNMENTS

```
RESULT 1
ADI28886
ID   ADI28886 standard; DNA; 5735 BP.
XX
AC   ADI28886;
XX
DT   22-APR-2004  (first entry)
XX
DE   Inducible, cardiac-preferred promoter MHCminTetO.
XX
KW   MHCminTetO; promoter; mouse; myosin; cardiomyopathy; cardiovascular-gen.;
KW   cardiant; antianginal; gene therapy; transgenic; ds.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   repeat_region   4282..4574
FT                   /*tag= b
FT                   /rpt_family= "DIRECT"
FT                   /note= "TetO binding site"
FT   repeat_unit     4282..4323
FT                   /*tag= a
XX
PN   WO2004005474-A2.
XX
PD   15-JAN-2004.
XX
PF   03-JUL-2003; 2003WO-US021035.
XX
PR   03-JUL-2002; 2002US-0393525P.
XX
PA   (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
PI   Robbins J;
XX
DR   WPI; 2004-091352/09.
XX
PT   New isolated nucleic acid molecule having inducible cardiac-preferred
PT   expression, useful for diagnosing or treating cardiac diseases, including
PT   ischemic heart disease, angina pectoris, myocardial infarction and
PT   endocarditis.
XX
PS   Claim 1; SEQ ID NO 1; 69pp; English.
```

```
XX
CC   The present sequence is that of MHCminTetO, an inducible, cardiac-
CC   preferred promoter sequence derived from the mouse alpha-myosin heavy
CC   chain promoter sequence. The native sequence was modified by insertion of
CC   a 7-repeat TetO binding site sequence. The promoter comprises a responder
CC   locus that is a copy number dependent, position independent locus in
CC   which various transgenes can be inserted. When uninduced, these
CC   transgenes are silent. When induced, the transgenes are very active.
CC   These genes can then be turned off using the inducible system. The
CC   promoter is useful for expressing operably linked sequences in a cardiac
CC   tissue-preferred expression pattern. Expression cassettes, host cells and
CC   transgenic animals are provided. The transgenic animals exhibit inducible
CC   cardiac-preferred expression of a nucleotide sequence of interest, e.g.
CC   ELC1a or glycogen synthase kinase 3-beta. These animals may have an
CC   altered susceptibility to cardiopathology and may be useful for
CC   identifying anti-cardiopathic compounds. The cardiopathology is
CC   especially a cardiomyopathy such as familial hypertrophic
CC   cardiomyopathies, dilated cardiomyopathies, peripartum cardiomyopathy,
CC   restrictive cardiomyopathies, ischaemic heart disease, angina pectoris,
CC   myocardial infarction, hypertensive heart disease and endocarditis
CC   (claimed).
XX
SQ   Sequence 5735 BP; 1458 A; 1455 C; 1545 G; 1277 T; 0 U; 0 Other;

Query Match             100.0%;  Score 5735;  DB 12;  Length 5735;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  5735;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGATCCTGCAAGTCACACAAGGTCTCCACCCACCAGGTGCCCTAGTCTCAATTCAGT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGATCCTGCAAGTCACACAAGGTCTCCACCCACCAGGTGCCCTAGTCTCAATTCAGT 60

Qy         61 TTCCATGCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTGGACTTTGTTTTTAT 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTCCATGCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTGGACTTTGTTTTTAT 120

Qy        121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGCTACCCAGCTCTAAGGGTGCCCGTGA 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGCTACCCAGCTCTAAGGGTGCCCGTGA 180

Qy        181 AGCCCTCAGAACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAGCAGAATAAAGCAATTT 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCCCTCAGAACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAGCAGAATAAAGCAATTT 240

Qy        241 TCCTTAAAGCCAAAATCTGCCCTAGACTCTTCTTCTGCCTCGGTCCCTGGGCTCT 300
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCCTTAAAGCCAAAATCTGCCCTAGACTCTTCTTCTGCCTCGGTCCCTGGGCTCT 300

Qy        301 AGGTGGGGAGGTGGGGCTTGGAAGAAGAAGTGGGAAGTGGGAAGCCGGATCCCTAG 360
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGGTGGGGAGGTGGGGCTTGGAAGAAGAAGTGGGAAGTGGGAAGCCGGATCCCTAG 360

Qy        361 GGCCCTGTGAAGTTCGGAGCCTTCCCGTACAGCACTGGCTCATAGATCCTCCTCAGCC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCCCTGTGAAGTTCGGAGCCTTCCCGTACAGCACTGGCTCATAGATCCTCCTCAGCC 420

Qy        421 AAACATAGCAAGAAGTGATACCCCTTTGTGACTTCCCAGTACCTGTCAGGTT 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAACATAGCAAGAAGTGATACCCCTTTGTGACTTCCCAGTACCTGTCAGGTT 480

Qy        481 GAAACAGGATTTAGAGAGCCTCTGAAGCTGAAGCTCGAACTCTGAAGCTCACCACCAAGCA 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GAAACAGGATTTAGAGAGCCTCTGAAGCTGAAGCTCGAACTCTGAAGCTCACCACCAAGCA 540

Qy        541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTACGCTGATAATATGCAGAGTGGGCCAC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTACGCTGATAATATGCAGAGTGGGCCAC 600

Qy        601 AGAAGTCCTGGGGTGTAGGAGTGAACTGACCCAGTGACTTTTCAGTCGGCAAAGGTATGACCCCC 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGAAGTCCTGGGGTGTAGGAGTGAACTGACCCAGTGACTTTTCAGTCGGCAAAGGTATGACCCCC 660

Qy        661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCATCCCAGGCAGGTCTCTAAGAGGACATG 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCATCCCAGGCAGGTCTCTAAGAGGACATG 720

Qy        721 GGATGAGAGATGAGTCATGTGGCATTCCAAACACAGCTATCCACCAGTGTCCCTTGCCCC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGATGAGAGATGAGTCATGTGGCATTCCAAACACAGCTATCCACCAGTGTCCCTTGCCCC 780

Qy        781 TTCCACTTAGCCAGGAGCAGTAACCTTAGCCCTAGCCCATCCTCCTCCCCATCCTCCCAG 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTCCACTTAGCCAGGAGCAGTAACCTTAGCCCTAGCCCATCCTCCTCCCCATCCTCCCAG 840

Qy        841 GACACACCCCTGTCTGCCAGTATTCATTTCTTCCTCACGTCCCTGTCCCTGTGACTTCCAT 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GACACACCCCTGTCTGCCAGTATTCATTTCTTCCTCACGTCCCTGTCCCTGTGACTTCCAT 900

Qy        901 TTGCAAGGCTTTGACCTCGCAGCTGCTGGAAGATAGAGTTGGCCCTAGGTGTGGCAA 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TTGCAAGGCTTTGACCTCGCAGCTGCTGGAAGATAGAGTTGGCCCTAGGTGTGGCAA 960

Qy        961 GCCATCTCAAGAGAAAGCAGACAACAGGGGACCAGATTTGGAAGGATCAGGAGAACTAAA 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCATCTCAAGAGAAAGCAGACAACAGGGGACCAGATTTGGAAGGATCAGGAGAACTAAA 1020

Qy       1021 TCACTGGGGGCCTGGGGTGAGAAAAAGAGTGAGTGAGTCCGCTCCAGTCAAGCCAAGC 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TCACTGGGGGCCTGGGGTGAGAAAAAGAGTGAGTGAGTCCGCTCCAGTCAAGCCAAGC 1080

Qy       1081 TAGTCCCCAGATACTCTGCCACCCAGGTCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TAGTCCCCAGATACTCTGCCACCCAGGTCTCGGGGTAGCTTTAGGAATGTGGGTC 1140

Qy       1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACCTACAGAC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACCTACAGAC 1200

Qy       1201 ACACTCGTGTGTGTGCTGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACCTGAGCTAG 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACACTCGTGTGTGTGCTGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACCTGAGCTAG 1260

Qy       1261 GCAACCAGAGCAATGGGCCCTGTGCTGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GCAACCAGAGCAATGGGCCCTGTGCTGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320

Qy       1321 CAGGGGAGGGAGAACAGAACAGAGAACAGAATAAGGAAGGAGAAAGGCCAGTCAATCAGA 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CAGGGGAGGGAGAACAGAACAGAGAACAGAATAAGGAAGGAGAAAGGCCAGTCAATCAGA 1380

Qy       1381 TGCAGTTCAGAAGGATGGGAAGCCAACACCAGCTTGAGCGAGAAGACAGAAAGGGAG 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TGCAGTTCAGAAGGATGGGAAGCCAACACCAGCTTGAGCGAGAAGACAGAAAGGGAG 1440

Qy       1441 AGATTCTGGCATAAGGAGGCCACAGAAAGAAGAGCCCCCCAAGTCTCCTCTTT 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGATTCTGGCATAAGGAGGCCACAGAAAGAAGAGCCCCCCAAGTCTCCTCTTT 1500

Qy       1501 ATACCCTCATCCCGTCCCAATTAAGCCCACTCTTCTCCTAGATCAGACCTGAGCTGC 1560
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATACCCTCATCCCGTCCCAATTAAGCCCACTCTTCTCCTAGATCAGACCTGAGCTGC 1560

Qy       1561 AGCGAAGAGACCCGTAGGAGCCTAAAAGAGCAAGACAAGAGATGTGAGAAGGGAG 1620
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 AGCGAAGAGACCCGTAGGAGCCTAAAAGAGCAAGACAAGAGATGTGAGAAGGGAG 1620

Qy       1621 GCAACCTAAGAGCCAGAGCCTAAAAGAGCTAAAAGGTGGCCAC 1680
              |||||||||||||||||||||||||||||||||||||||||||
Db       1621 GCAACCTAAGAGCCAGAGCCTAAAAGAGCTAAAAGGTGGCCAC 1680

Qy       1681 CTGTGCACACAGAGGGTCGAGGACTGGTGTGAGAGCCTCAAGATAAGGATGATGCTCAGA 1740
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTGTGCACACAGAGGGTCGAGGACTGGTGTGAGAGCCTCAAGATAAGGATGATGCTCAGA 1740

Qy       1741 TCAGCAGATGTAGTAATGTCCCCTTAGATCCATCCCAGGCAGGTCTCTAAGAGGACATG 1800
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATGGGCGGGGGGGATTCTGGGGGGGAGGAGAAGGAGCCTGAGAACAG 1800
```

```
Db  1741 ATGGGCGGGGGGGGGATTCTGGGGGGGAGAGAGAGAGGTGAGAGAGAGCCTGGAACAG 1800

Qy  1801 AGAATCTGAAGCGCTGAAACGATACCATAAAGGGAAGAACACCAGGCTACCTTTAGATG 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGAATCTGAAGCGCTGAAACGATACCATAAAGGGAAGAACACCAGGCTACCTTTAGATG 1860

Qy  1861 TAAATCATGAAAGACAGGAGAGGAGAAGCTGAGAGAGTAGAGAAGGACCCCGGGGCCAAGA 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TAAATCATGAAAGACAGGAGAGGAGAAGCTGAGAGAGTAGAGAAGGACCCCGGGGCCAAGA 1920

Qy  1921 CATGGAAGCAAGGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGTGAAGGCAGAG 1980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CATGGAAGCAAGGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGTGAAGGCAGAG 1980

Qy  1981 CCCTGGTATGAGCACCAGAACAGCAGAGCTAGGGTTAATGTGAGACAGGAACAGAGAAG 2040
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 CCCTGGTATGAGCACCAGAACAGCAGAGCTAGGGTTAATGTGAGACAGGAACAGAGAAG 2040

Qy  2041 GTAGACACAGACAGACAGAGACGGGGAGCCAGTGAGAGCCAGAGCAAGGGGGCCTAGGA 2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 GTAGACACAGACAGACAGAGACGGGGAGCCAGTGAGAGCCAGAGCAAGGGGGCCTAGGA 2100

Qy  2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGC 2160

Qy  2161 TGGCTTGGGAGGCCAGCTGGAAAGAGTATGTGAGAGCCACACAAGGGACAAGGGGCTAGGA 2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TGGCTTGGGAGGCCAGCTGGAAAGAGTATGTGAGAGCCACACAAGGGACAAGGGGCTAGGA 2220

Qy  2221 AAGGAAGAGAGAGGGCAAACCAGGCCACACAAGGGGCCAGAGCCCAGAACTGAGTTAACTC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AAGGAAGAGAGAGGGCAAACCAGGCCACACAAGGGGCCAGAGCCCAGAACTGAGTTAACTC 2280

Qy  2281 CTTCCTTGTTGCATCTTCCATAGAGGCAGTGGAACTCTGTGACCACCATCCCCATGA 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CTTCCTTGTTGCATCTTCCATAGAGGCAGTGGAACTCTGTGACCACCATCCCCATGA 2340

Qy  2341 GCCCCCACTACCCATACCAAGTTTGGCCTGAGTTGGCCATTCTAGGTTCCCTGAGGACAGAG 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GCCCCCACTACCCATACCAAGTTTGGCCTGAGTTGGCCATTCTAGGTTCCCTGAGGACAGAG 2400

Qy  2401 CCTGGCCTTTGTCTCTTGACCTGACCCAAGCTGACATGTTCTCAGTACCTTTGAAT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CCTGGCCTTTGTCTCTTGACCTGACCCAAGCTGACATGTTCTCAGTACCTTTGAAT 2460

Qy  2461 GCCCTCAAGAGCTTGAGAACCAGCAGTGACATATTAGGCCATGGCATGGCTAACCCTGGAGCT 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 GCCCTCAAGAGCTTGAGAACCAGCAGTGACATATTAGGCCATGGCATGGCTAACCCTGGAGCT 2520

Qy  2521 TGCACACAGGAGCCTCAAGGTGACCTCAGGGACACAGTGGCCCTTTATCC 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 TGCACACAGGAGCCTCAAGGTGACCTCAGGGACACAGTGGCCCTTTATCC 2580

Qy  2581 CCAAAGAGCAACCATTTGCATAGGTGGCTGCATAGGTGGCTGCGCAAATGGGAATGCAAGGTTGAATCAGGTC 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 CCAAAGAGCAACCATTTGCATAGGTGGCTGCATAGGTGGCTGCGCAAATGGGAATGCAAGGTTGAATCAGGTC 2640

Qy  2641 CCTTCAAGAATACTGCAAGAGAGACCTTAAGAGAGGGGTATGCTCCTGCCC 2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 CCTTCAAGAATACTGCAAGAGAGACCTTAAGAGAGGGGTATGCTCCTGCCC 2700

Qy  2701 CCACCCACCATAAGGGAGTGAACTATCCTAGGGGGCTGGCGACCTGGGAGACACCAC 2760
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 CCACCCACCATAAGGGAGTGAACTATCCTAGGGGGCTGGCGACCTGGGAGACACCAC 2760

Qy  2761 ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCCGCCCGTGTCCTGCCACCTCCCACAC 2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCCGCCCGTGTCCTGCCACCTCCCACAC 2820

Qy  2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGAGATAGAGGGTGGTAGCAGGGAGAGTGT 2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGAGATAGAGGGTGGTAGCAGGGAGAGTGT 2880
```

```
Qy  2881 TCCTCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGCTTTGTGTCTCCTGAA 2940
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 TCCTCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGCTTTGTGTCTCCTGAA 2940

Qy  2941 CACAATGTCTACTTAGTTATTATACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC 3000
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 CACAATGTCTACTTAGTTATTATACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC 3000

Qy  3001 TGAAAAGACAGCAGCCCTGGAGGACAGGGTGTCTCTGAAGCCTTGGGTGCTTGGGTGATGGTG 3060
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TGAAAAGACAGCAGCCCTGGAGGACAGGGTGTCTCTGAAGCCTTGGGTGCTTGGGTGATGGTG 3060

Qy  3061 CCACAAAGGAGGGCATGAGTGTGAGTAAGGGCCCCTATCCATGGAATCTGGAGCCTGGGCCCAACTGGTG 3120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 CCACAAAGGAGGGCATGAGTGTGAGTAAGGGCCCCTATCCATGGAATCTGGAGCCTGGGCCCAACTGGTG 3120

Qy  3121 GGAAGGGGTCAGTCTGCAGAGCCCTGCCAGGCAATTCAAAGCACCTGCATCCTCTGGGCAGCCTGGGGA 3180
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 GGAAGGGGTCAGTCTGCAGAGCCCTGCCAGGCAATTCAAAGCACCTGCATCCTCTGGGCAGCCTGGGGA 3180

Qy  3181 TAAATCTCTGGCCTGCCAGGCATTCAAAGCACCCTTTCTCCCCTCAGGAGTGGAGGTTGCAGAGGAGGGTAAAAA 3240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 TAAATCTCTGGCCTGCCAGGCATTCAAAGCACCCTTTCTCCCCTCAGGAGTGGAGGTTGCAGAGGAGGGTAAAAA 3240

Qy  3241 GGCGAAGGGAGCAACCCCCCTTCCCACCTCCCATCAGGAGTGGCGATATATGGTCAGTAGTGATCAGTAGAGGCAAGAAA 3300
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GGCGAAGGGAGCAACCCCCCTTCCCACCTCCCATCAGGAGTGGCGATATATGGTCAGTAGTGATCAGTAGAGGCAAGAAA 3300

Qy  3301 CTGGCCTTCCCCCTTCCCCATCCCCACCTCCCATCAGGAGTGGCTAACTCCAGCTGTTCAGTTCTGTGCGCTGAGGTTACCAGAACTGCA 3360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CTGGCCTTCCCCCTTCCCCATCCCCACCTCCCATCAGGAGTGGCTAACTCCAGCTGTTCAGTTCTGTGCGCTGAGGTTACCAGAACTGCA 3360

Qy  3361 CCTACACATGTCAAACATCATGGTCAACATCATGGTCAAATGGGTAAATGTGTAAAGTCTGTGTGCATGTGTGTCTG 3420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 CCTACACATGTCAAACATCATGGTCAACATCATGGTCAAATGGGTAAATGTGTAAAGTCTGTGTGCATGTGTGTCTG 3420

Qy  3421 GGAAATCTGCAGGCTTAACTGGGTTAACTGGGTTACCAGACTCTTGTCGTCGTTCAGTTCTGTGCCTGAGGTTACCAGAACTGCA 3480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GGAAATCTGCAGGCTTAACTGGGTTAACTGGGTTACCAGACTCTTGTCGTCGTTCAGTTCTGTGCCTGAGGTTACCAGAACTGCA 3480

Qy  3481 ACTGAAAACCGGGCACATGGCCATGGCATGGCATGGCATGCAGTCTTTGGAGGACACCTGTCTTTGGAGGACACCTGTCTTCAGATGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 ACTGAAAACCGGGCACATGGCCATGGCATGGCATGGCATGCAGTCTTTGGAGGACACCTGTCTTCAGATGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600

Qy  3601 TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT 3660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601 TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT 3660

Qy  3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGGATGGGCAACAGAGAC 3720
         |||||||||||||||||||||||||||||||||||||||||||
Db  3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGGATGGGCAACAGAGAC 3720

Qy  3721 TACAGACAGTATCTGTACAGAGTAAGGACGAGAGGAGGGGTGTAGAATTCTCTTACTA 3780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 TACAGACAGTATCTGTACAGAGTAAGGACGAGAGGAGGGGTGTAGAATTCTCTTACTA 3780

Qy  3781 TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGCTCCCTGCAAAGTGCTCCCTAGACATCATGACTT 3840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGCTCCCTAGACATCATGACTT 3840

Qy  3841 TGTCTCTGGGGAGCCAGCACTGTGAACTTCAGGTCTGAGAGTCAACAGAGAC 3900
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTCTCTGGGGAGCCAGCACTGTGAACTTCAGGTCTGAGAGTCAACAGAGAC 3900

Qy  3901 GCCTGAAGCTATGCAGATAGCCAGGTTGAAAGGGGAAGGGGAAGGGGAGAGGCCTGGGATGGGAGC 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 GCCTGAAGCTATGCAGATAGCCAGGTTGAAAGGGGAAGGGGAGAGGCCTGGGATGGGAGC 3960
```

```
Qy  3961 TTGTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAGAGAGAGGTGACCCTTACCCAG 4020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 TTGTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAGAGAGAGGTGACCCTTACCCAG 4020

Qy  4021 TTGTTCAACTCACCCTTCAGATTAAAATAACTGAGGTAAGGGCCTGGGTAGGGGAGGTG 4080
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 TTGTTCAACTCACCCTTCAGATTAAAATAACTGAGGTAAGGGCCTGGGTAGGGGAGGTG 4080

Qy  4081 GTGTGAGACGCTCCTGTCTCCTCTGCATGCCCGAGGGCCCTTGGGGAGGAGGAATGT 4140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTGTGAGACGCTCCTGTCTCCTCTGCATGCCCGAGGGCCCTTGGGGAGGAGGAATGT 4140

Qy  4141 GCCCAAGGACTAAAAAAAGGCCATGGAGCCAGAGGGCGCAGAGGGCAACAGACCTTTCATGG 4200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 GCCCAAGGACTAAAAAAAGGCCATGGAGCCAGAGGGCGCAGAGGGCAACAGACCTTTCATGG 4200

Qy  4201 GCAAACCTTGGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAATCGATACCCTTCTTC 4260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201 GCAAACCTTGGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAATCGATACCCTTCTTC 4260

Qy  4261 TTTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGA 4320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4261 TTTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGA 4320

Qy  4321 AAGTGCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTC 4380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4321 AAGTGCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTC 4380

Qy  4381 CCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAA 4440
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4381 CCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAA 4440

Qy  4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA 4500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA 4500

Qy  4501 CCACTCCCTATCAGTGATAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGAT 4560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4501 CCACTCCCTATCAGTGATAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGAT 4560

Qy  4561 AGAGAAAGTGAAAGTCGAGTCGGTACCAGCAGAGCACTGACCAGAATTTAGGCAGCAGGCA 4620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4561 AGAGAAAGTGAAAGTCGAGTCGGTACCAGCAGAGCACTGACCAGAATTTAGGCAGCAGGCA 4620

Qy  4621 TATGGGGATGGGATATAAAGGGGCTGGAGCACTGACCTGTCAGAGATTTCTCCAACCCA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4621 TATGGGGATGGGATATAAAGGGGCTGGAGCACTGACCTGTCAGAGATTTCTCCAACCCA 4680

Qy  4681 GGTAAGAGGAGGAGTTTCGGGTGGGGGGCTCTCACCCACCACCAGACCTCTCCCCACCTAGAA 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4681 GGTAAGAGGAGGAGTTTCGGGTGGGGGGCTCTCACCCACCACCAGACCTCTCCCCACCTAGAA 4740

Qy  4741 GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCGGGTTCAGCGGGTCTGACAGGGTGGCCTT 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4741 GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCGGGTTCAGCGGGTCTGACAGGGTGGCCTT 4800

Qy  4801 CCACCAGCCTGGGAAGTTCTCAGTGCAGGAGGAGTTTCCACAGAAAACACTGATGATGCCCT 4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4801 CCACCAGCCTGGGAAGTTCTCAGTGCAGGAGGAGTTTCCACAGAAAACACTGATGATGCCCT 4860

Qy  4861 TCCCTTACGCGTCTTCTTCCATCTTCCATCTTCCTGGGATGCTCCCCCGTCTTGGTTTATC 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4861 TCCCTTACGCGTCTTCTTCCATCTTCCATCTTCCTGGGATGCTCCCCCGTCTTGGTTTATC 4920

Qy  4921 TTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGTTCCTTCTCCACTCTTTCTCTACTGT 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4921 TTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGTTCCTTCTCCACTCTTTCTCTACTGT 4980

Qy  4981 CTCCGTGTCCTTGCCTTGCCTTGCGTGTCTTGCGTGTCCTTCCTTCACCCCATTTCTCACTTCACC 5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4981 CTCCGTGTCCTTGCCTTGCCTTGCGTGTCTTGCGTGTCCTTCCTTCACCCCATTTCTCACTTCACC 5040

Qy  5041 TTTTCTCCCCTTCTCATTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5041 TTTTCTCCCCTTCTCATTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100

Qy  5101 CTTCCTTCCTTCCTTCCTTCCTTC 5160
         |||||||||||||||||||||||
Db  5101 CTTCCTTCCTTCCTTCCTTCCTTC 5160

Qy  5161 CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTCCCACCCTTATGTGAACAATCTTCCA 5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5161 CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTCCCACCCTTATGTGAACAATCTTCCA 5220

Qy  5221 GTGAGCCACAGCTTCAGTGCTGCTGCGGTGCTCTCTCTAACCTTCCTCACCCCTCGGCTTGTC 5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5221 GTGAGCCACAGCTTCAGTGCTGCTGCGGTGCTCTCTCTAACCTTCCTCACCCCTCGGCTTGTC 5280

Qy  5281 CTGTTCCATCCTGTCAGATCTCTAGATTGGTCTCCAGCGCGCCACTGTGGTGGCTCGTGTGTCCAC 5340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5281 CTGTTCCATCCTGTCAGATCTCTAGATTGGTCTCCAGCGCGCCACTGTGGTGGCTCGTGTGTCCAC 5340

Qy  5341 GCCTGTTCTGTCTCTCTGTCGTGTCCAGCTCCTGAAAAGTAACCAGGTGAGAATGTTTCCCCTGGC 5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5341 GCCTGTTCTGTCTCTCTGTCGTGTCCAGCTCCTGAAAAGTAACCAGGTGAGAATGTTTCCCCTGGC 5400

Qy  5401 AATTCTTCAGGATTCTCTGAAAAGTAACCAGGTGAGAATGTTTCCCCTGCTGCCCAGCTGCCTGCCGGCA 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5401 AATTCTTCAGGATTCTCTGAAAAGTAACCAGGTGAGAATGTTTCCCCTGCTGCCCAGCTGCCTGCCGGCA 5460

Qy  5461 TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTCTTCTGCCCAGCTGCCCGGCA 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5461 TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTCTTCTGCCCAGCTGCCCGGCA 5520

Qy  5521 CTCTTAGCAAACCTCAGGCACCCTACCCCACCACATAGACCTCTGACACAGAACCTCTGACCACT 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5521 CTCTTAGCAAACCTCAGGCACCCTACCCCACCACATAGACCTCTGACACAGAACCTCTGACCACT 5580

Qy  5581 TTACATGGAGTCCTGGTCGGTGGGAGCCATAGGCTACGGTGTAAAAGAGGCAGGAGGAAGTGGT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5581 TTACATGGAGTCCTGGTCGGTGGGAGCCATAGGCTACGGTGTAAAAGAGGCAGGAGGAAGTGGT 5640

Qy  5641 GGTGTAGGAAAGTCAGGACTTCACAGAAGTCAGGAGCTTCACCCACCCAGACCACCCACCAGACACAGA 5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5641 GGTGTAGGAAAGTCAGGACTTCACAGAAGTCAGGAGCTTCACCCACCCAGACCACCCACCAGACACAGA 5700

Qy  5701 TCCCTCCTATCTCCCCATAAGAGTTTGGAGTCGAC 5735
         |||||||||||||||||||||||||||||||||||
Db  5701 TCCCTCCTATCTCCCCATAAGAGTTTGGAGTCGAC 5735
```

XX
PI Field LJ, Pasumarthi KBS;
XX
DR WPI; 2001-102575/11.
XX
PT Increasing proliferative potential of cardiomyocyte cell which is used
PT for screening activity of biological or pharmacological agent, involves
PT increasing the level of cyclin D2 activity in cardiomyocyte cell.
XX
PS Example 1; Page 62-66; 67pp; English.
XX
CC The present sequence is mouse alpha-cardiac myosin heavy chain (MHC)
CC promoter. This sequence is used in the preparation of a MHC-CYCD2 fusion
CC gene. The cyclin D2 (CYCD2) cDNA sequence is used to increase the
CC proliferative potential of a cardiomyocyte cell by increasing the level
CC of cyclin D2 in the cardiomyocyte cell. The transgenic animals
CC expressing cyclin D2 have sustained atrial and ventricular cardiomyocyte
CC DNA synthesis. Cardiomyocyte cells with enhanced proliferative potential
CC are useful for screening the activity of biological or pharmacological
CC agents on cardiomyocyte cells. Genetically modified cardiomyocyte cells
CC are useful for delivering therapeutics to mammals. The cells are also
CC useful to target an improvement of the contractile function of the heart
CC of the patient, for e.g. in the treatment of contractile losses due to
XX infarcts or cardiomyopathies
SQ Sequence 5443 BP; .1352 A; 1399 C; 1492 G; 1200 T; 0 U; 0 Other;

Query Match   88.0%; Score 5045; DB 5; Length 5443;
Best Local Similarity   93.9%; Pred. No. 0;
Matches 5383; Mismatches   0; Conservative   60; Indels 292; Gaps   1;

Qy   1 GGATCCTGCAAGGTCACACAAGGTCTCCACCACCAGGTGCCCAGTCTCAATTCAGT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 GGATCCTGCAAGGTCACACAAGGTCTCCACCACCAGGTGCCCAGTCTCAATTCAGT   60

Qy   61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCAGAGCTAATTTGGACTTTGTTTTAT 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCAGAGCTAATTTGGACTTTGTTTTAT 120

Qy   121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGCTACCCAGTCTCTAAGGGTGCCCGTGA 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGCTACCCAGTCTCTAAGGGTGCCCGTGA 180

Qy   181 AGCCCTCAGACCTTGGAGCCTTTGCAACAGCCCCTTAGGTGGAAGCAGAATAAAGCAATTT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 AGCCCTCAGACCTTGGAGCCTTTGCAACAGCCCCTTAGGTGGAAGCAGAATAAAGCAATTT 240

Qy   241 TCCTTAAAGCCAAAATCCTGCCTCTTCTTCTCGAGACTCTCTTCTGCCTCCGGTCCCTGGGCTCT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TCCTTAAAGCCAAAATCCTGCCTCTTCTTCTCGAGACTCTCTTCTGCCTCCGGTCCCTGGGCTCT 300

Qy   301 AGGGTGGGGAGGGGCTTGGAAGAAGAGAGGTGGGAGTGGCAAAAGCCGATCCCTAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 AGGGTGGGGAGGGGCTTGGAAGAAGAGAGGTGGGAGTGGCAAAAGCCGATCCCTAG 360

Qy   361 GGCCCTGTGAAGTTCGGAAGCTTCCCTGTACAGCACTGCTTCATAGATCCTCCTCCAGCC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGCCCTGTGAAGTTCGGAAGCTTCCCTGTACAGCACTGCTTCATAGATCCTCCTCCAGCC 420

Qy   421 AAACATAGCAAGAAGTGATACCTTCCTTTGTGACTTCCCCAGGGCCCAGTTACCTGTCAGGTT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 AAACATAGCAAGAAGTGATACCTTCCTTTGTGACTTCCCCAGGGCCCAGTTACCTGTCAGGTT 480

Qy   481 GAAACAGGATTTAGAGAGCCTCTGAACTCTGAACTCTGAAGCTCATCCACCACAAGCA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAAACAGGATTTAGAGAGCCTCTGAACTCTGAACTCTGAAGCTCATCCACCACAAGCA 540

Qy   541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTACGCTTAGTGAGAGCTGGGCCAC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTACGCTTAGTGAGAGCTGGGCCAC 600

Qy   601 AGAAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGGTATGACCCCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGAAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGGTATGACCCCC 660

Qy   661 TCAGCAGAGATGAGTAATGTCCCCTTAGATCCCATCCCAGCAGGTCTCTAAGAGGACATG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCAGCAGAGATGAGTAATGTCCCCTTAGATCCCATCCCAGCAGGTCTCTAAGAGGACATG 720

Qy   721 GGATGAGAGATGAGTAGTCATGTGGCATTCCAAACACAGCTATCCACAGTGTCCCTTGCCCC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GGATGAGAGATGAGTAGTCATGTGGCATTCCAAACACAGCTATCCACAGTGTCCCTTGCCCC 780

Qy   781 TTCCACTTAGCCAGGAGGACGTAACCTTAGCCTATCCTTTCCTCCCCATCCTCCCAG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TTCCACTTAGCCAGGAGGACGTAACCTTAGCCTATCCTTTCCTCCCCATCCTCCCAG 840

Qy   841 GACACACCCCTGGTCTGCAGTATTCATTCTTCCTTCACGTCCCCTGTGACTTCCAT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACACACCCCTGGTCTGCAGTATTCATTCTTCCTTCACGTCCCCTGTGACTTCCAT 900

Qy   901 TTGCAAGGCTTTGACCTCTGCAGTGCTGGAAGATAGAGTTGGCCTAGGTGTGGCAA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTGCAAGGCTTTGACCTCTGCAGTGCTGGAAGATAGAGTTGGCCTAGGTGTGGCAA 960

Qy   961 GCCATCTCAAGAGAAAAGCAGACAACAGGGGGACCCAGATTTTGGAAGGATCAGGAACTAAA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GCCATCTCAAGAGAAAAGCAGACAACAGGGGGACCCAGATTTTGGAAGGATCAGGAACTAAA 1020

Qy   1021 TCACTGGCGGGCCTTGGGGGTAGAAAAAAGAGTGAGTCCGCTCCAGCTAAGCCAAGC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCACTGGCGGGCCTTGGGGGTAGAAAAAAGAGTGAGTCGAGTCCGCTCCAGCTAAGCCAAGC 1080

Qy   1081 TAGTCCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAGTCCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140

Qy   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACCTACAGAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACCTACAGAC 1200

Qy   1201 ACACTCGTGTGTGGCCAGACTCCTGTTCAACAGCCCTGTTCTGACCACTGAGCTAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACACTCGTGTGTGGCCAGACTCCTGTTCAACAGCCCTGTTCTGACCACTGAGCTAG 1260

Qy   1261 GCAACCAGACATGGCCCTTGTGCTTGAGGATGAACAATAGCAAAAACAG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCAACCAGACATGGCCCTTGTGCTTGAGGATGAACAATAGCAAAAACAG 1320

Qy   1321 CAGGGGAGGGGAGAACAGAGAACAGAATATAAGGACGAAGGCCAGTCAATCAGA 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CAGGGGAGGGGAGAACAGAGAACAGAATATAAGGAAGGAAGGCCAGTCAATCAGA 1380

Qy   1381 TGCAGTCAGAGAGATGGGAAGCCAACACACAGCTTGAGCAGAGAAACAGAAAGGGAG 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCAGTCAGAGAGATGGGAAGCCAACACACAGCTTGAGCAGAGAAACAGAAAGGGAG 1440

Qy   1441 AGATTCTGGGCATAAGGAGGAGCCACAGAAGAGAGAAGAGCCCCCAAGTCTCCTCTTT 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AGATTCTGGGCATAAGGAGGAGCCACAGAAGAGAGAAGAGCCCCCAAGTCTCCTCTTT 1500

Qy   1501 ATACCCTCATCCCGTCTCCCAATTAAGCCACTTCTTCCTAGATCAGACCTGAGCTGC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATACCCTCATCCCGTCTCCCAATTAAGCCACTTCTTCCTAGATCAGACCTGAGCTGC 1560

Qy   1561 AGCGAAGAGAGACCCGTAGGGAGGAGATGTGTGGAGAAGTCCAGG 1620
            |||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGCGAAGAGAGACCCGTAGGGAGGAGATGTGTGGAGAAGTCCAGG 1620

Qy   1621 GCAACCTAAGAGCCAGAGCCTAAAAGAGCAAGAGAGCTGCTTCAAAGGTGGCCAGG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCAACCTAAGAGCCAGAGCCTAAAAGAGCAAGAGAGATAAGGTGCTTCAAAGGTGGCCAGG 1680

Qy   1681 CTGTGCACACAGAGGGTCGAGGAGGGTGTGGAGCCTCGAGACGTATGAGCTCAGA 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CTGTGCACACAGAGGGTCGAGGAGGGTGTGGAGCCTCGAGACGTATGAGCTCAGA 1740

```
Qy  2821  TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGTGGGAGCTGGTGAGCTGGTGAGCAGGAGGAGTGT  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGTGGGAGCTGGTGAGCTGGTGAGCAGGAGGAGTGT  2880

Qy  2881  TCCTGGGTGTGAGGTGTAGGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGTCTTGTCTCCTGAA  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  TCCTGGGTGTGAGGTGTAGGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGTCTTGTCTCCTGAA  2940

Qy  2941  CACAATGTCTACTTAGTTATAACACGGCATGACCTGCTAAAGACCCAACATCTACGACCTC  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  CACAATGTCTACTTAGTTATAACACGGCATGACCTGCTAAAGACCCAACATCTACGACCTC  3000

Qy  3001  TGAAAAGACAGCAGCCCTGGAGACACAGGGGTTGTCTCTGAGCCTTGGTGCTTGATGGTG  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TGAAAAGACAGCAGCCCTGGAGACACAGGGGTTGTCTCTGAGCCTTGGTGCTTGATGGTG  3060

Qy  3061  CCACAAGGAGGGCATGAGTGTGAGTATAGGCCATGAAGGGCACTTG  3120
          |||||||||||||||||||||||||||||||||||||||||||||
Db  3061  CCACAAGGAGGGCATGAGTGTGAGTATAGGCCATGAAGGGCACTTG  3120

Qy  3121  GGAAGGGGTCAGTCTGCGAGAGCCCTATCCGAGCAGCACCTGGAATCTGGAGCCTGGGGA  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  GGAAGGGGTCAGTCTGCGAGAGCCCTATCCGAGCAGCACCTGGAATCTGGAGCCTGGGGA  3180

Qy  3181  TAAATCTCTGGGCCTGCCAGGCATTCAAAGCAGCACCTGCATCCTCCTGGCCAGCCTGGGGA  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  TAAATCTCTGGGCCTGCCAGGCATTCAAAGCAGCACCTGCATCCTCCTGGCCAGCCTGGGGA  3240

Qy  3241  GGCGGAAGGGAGCAACCCCCACTTATACCCTTTCTCCCCAGCCCCAGATTAACACCT  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GGCGGAAGGGAGCAACCCCCACTTATACCCTTTCTCCCCAGCCCCAGATTAACACCT  3300

Qy  3301  CTGGCCTTCCCCCTTCCCCCTTCCCACCTCCCACTCCAGGAGGTTGCAGAGGGTAAAAA  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CTGGCCTTCCCCCTTCCCCCTTCCCACCTCCCACTCCAGGAGGTTGCAGAGGGTAAAAA  3360

Qy  3361  CCTACATGTCAAACATCATGTCACGATATATGGATCAGTATGTAGAGGCAAGAAA  3420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  CCTACATGTCAAACATCATGTCACGATATATGGATCAGTATGTAGAGGCAAGAAA  3420

Qy  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGTGTGTCTG  3480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGTGTGTCTG  3480

Qy  3481  ACTGAAAACGGGACATGGCGTGCGAGCTGTCAGTTCTGTGCTCGTGAGGCTCAGACTGCA  3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  ACTGAAAACGGGACATGGCGTGCGAGCTGTCAGTTCTGTGCTCGTGAGGCTCAGACTGCA  3540

Qy  3541  GGTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCTTCCATGGTTTAAAGAGAT  3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GGTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCTTCCATGGTTTAAAGAGAT  3600

Qy  3601  TGGATGATGGCTGCATCTGCATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT  3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  TGGATGATGGCTGCATCTGCATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT  3660

Qy  3661  CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGCAACAGAGAC  3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGCAACAGAGAC  3720

Qy  3721  TACAGACAGTATCTGTACAGGTAAGGAGGAGGAGGAGGGGGTGTAGAATTCTCTTACTA  3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  TACAGACAGTATCTGTACAGGTAAGGAGGAGGAGGAGGGGGTGTAGAATTCTCTTACTA  3780

Qy  3781  TCAAAGGGAAACTGAGTGTGCACCTGCAAAGTGGATGCTGTCTCCTAGACATCATGACTT  3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  TCAAAGGGAAACTGAGTGTGCACCTGCAAAGTGGATGCTGTCTCCTAGACATCATGACTT  3840

Qy  3841  TGTCTCTGGGAGCCAGCACTGTGGAACTTCAGGTCTGTGGAACTTCAGGTTGAAAGGGGTTGAAAGGGGTTGAAAGGGAGAGCCTCGGGATGGGAGC  3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTCTCTGGGAGCCAGCACTGTGGAACTTCAGGTCTGTGAGAGACTGTGGAACTTCAGGTTTAGAGAGTAGGAGGCCTCGGGATGGGAGC  3900

Qy  3901  GCCTGAAGCTATGCGAGATAGCCAGGGTTGAAAGGGGGAGAGCCTCGGGATGGGAGC  3960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  GCCTGAAGCTATGCGAGATAGCCAGGGTTGAAAGGGGGAGAGCCTCGGGATGGGAGC  3960
```

```
Qy  1741  ATGGGGCCGGGGGGGGGGATTCTCTGGGGGGGAGAGAAGGTGAGAGGAGCCTGGAACAG  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  ATGGGGCCGGGGGGGGGGATTCTCTGGGGGGGAGAGAAGGTGAGAGGAGCCTGGAACAG  1800

Qy  1801  AGAATCTGGAAGCGCTGGAAACGACGATACCATAAAGGGAGAACCCAGGCTACCTTTAGATG  1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGAATCTGGAAGCGCTGGAAACGACGATACCATAAAGGGAGAACCCAGGCTACCTTTAGATG  1860

Qy  1861  TAAATCATGAAAGACAGGGAGGAAGCTGGAGAGTAGAAGGACCCCGGGGCAAGA  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  TAAATCATGAAAGACAGGGAGGAAGCTGGAGAGTAGAAGGACCCCGGGGCAAGA  1920

Qy  1921  CATGGAAGCAAGACAAGCCAGTTGAGCGCTCCGTGAAATCAGCGTTGAGAGGCAGAG  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  CATGGAAGCAAGACAAGCCAGTTGAGCGCTCCGTGAAATCAGCGTTGAGAGGCAGAG  1980

Qy  1981  CCCTGGTATGAGCACAGAACACAGAGACCTAGGGTTAATGTGAGACAGGAACAGAAG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  CCCTGGTATGAGCACAGAACACAGAGACCTAGGGTTAATGTGAGACAGGAACAGAAG  2040

Qy  2041  GTAGACACAGGAACAGACAGAGCGGGGACGCGGGGACGCGGGGTAACAAAGGAATGGTCCTTCTCAC  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GTAGACACAGGAACAGACAGAGCGGGGACGCGGGGACGCGGGGTAACAAAGGAATGGTCCTTCTCAC  2100

Qy  2101  CTGTGGCCAGAGCGTCCATCTGTCTCCACATATCTAGAGATGTTCATCAGACTGCAGGGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  CTGTGGCCAGAGCGTCCATCTGTCTCCACATATCTAGAGATGTTCATCAGACTGCAGGGC  2160

Qy  2161  TGGCTTGGGAGGCAGCTGGAAAGAGTATGTGAGAGCAAGGGGCCTAGGA  2220
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TGGCTTGGGAGGCAGCTGGAAAGAGTATGTGAGAGCAAGGGGCCTAGGA  2220

Qy  2221  AAGGAAGAGAGGGCAAACCAGGCCACACAGAGGGCAGAGCCCAGAACTGAGTTAACTC  2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  AAGGAAGAGAGGGCAAACCAGGCCACACAGAGGGCAGAGCCCAGAACTGAGTTAACTC  2280

Qy  2281  CTTCCTTGTTGCATCTTCCATAGGAGCCAGTGGAACTCTGTGACCACCATGA  2340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CTTCCTTGTTGCATCTTCCATAGGAGCCAGTGGAACTCTGTGACCACCATGA  2340

Qy  2341  GCCCCCACTACCCATCCAAGTTTGGCCTGAGTGGCATTCTAGGTTCCCTGAGGACAGAG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GCCCCCACTACCCATCCAAGTTTGGCCTGAGTGGCATTCTAGGTTCCCTGAGGACAGAG  2400

Qy  2401  CCTGGGCCTTTGTCTTGACCCAAGCTGACCCAAGTTCTCAGTGTTCTCAGTTGAAT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||     |
Db  2401  CCTGGGCCTTTGTCTTGACCCAAGCTGACCCAAGTTCTCAGTTTCTCAGTTATCAT  2460

Qy  2461  GCCCTCAAGAGCTTGAGAACCAGGCAGTGACATATTAGGCCATGGCTAACCCTGGAGCT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  GCCCTCAAGAGCTTGAGAACCAGGCAGTGACATATTAGGCCATGGCTAACCCTGGGACT  2520

Qy  2521  TGCACACAGGAGCCTCAAGTGACCTCCAGAGCTGCAGCAGTGACCTGGCCTTTATCC  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  TGCACACAGGAGCCTCAAGTGACCTCCAGAGCACACAGTGCCTGGCCTTTATCC  2580

Qy  2581  CCAAAGAGCAACCATTTGCATAGGTGCTGCAAATGGAATCAGGTTGAATCAGTC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  CCAAAGAGCAACCATTTGCATAGGTGCTGCAAATGGAATCAGGTTGAATCAGTC  2640

Qy  2641  CCTTCAAGAATACTGCATGCAAGACCTTAAGACCCCCTGAGAGAGGGGTATGCCTGCCC  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  CCTTCAAGAATACTGCATGCAAGACCTTAAGACCCCCTGAGAGAGGGGTATGCCTGCCC  2700

Qy  2701  CCACCCCACCTAGGGGTGAACTATCCTAGGGGGTGGCGACCTTGGGGGAGACACCAC  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  CCACCCCACCTAGGGGTGAACTATCCTAGGGGGTGGCGACCTTGGGGGAGACACCAC  2760

Qy  2761  ATTACTGGAGTGCTGAGCCCCAGAAAACTGACCCCACCCTGTCCTGCCCACCTCCCACAC  2820
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  ATTACTGGAGTGCTGAGCCCCAGAAAACTGACCCCACCCTGTCCTGCCCACCTCCCACAC  2820
```

```
Db  4689 CTCCGTGCCTGCCTGCCTTCTTGCGTGTCCTTCCTTCCACCCATTTCTCACTTCACC 4748
Qy  5041 TTTTCTCCCCTTCCATTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100
Db  4749 TTTTCTCCCCTTCCATTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4808
Qy  5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5160
Db  4809 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4868
Qy  5161 CTGTTCAGAGTGTGAGAATCACACCTGGGTTCCACACCCTTATGTAAACAATCTTCCA 5220
Db  4869 CTGTGTCAGAGTGTGAGAATCACACCTGGGTTCCACACCCTTATGTAAACAATCTTCCA 4928
Qy  5221 GTGAGCCACAGCTTCAGTGCTGCTGGCTCTGGGTGCTCTCTGGTGCTGTGTCCAC 5280
Db  4929 GTGAGCCACAGCTTCAGTGCTGCTGGCTCTGGGTGCTCTCTGGTGCTGTGTCCAC 4988
Qy  5281 CTGTTCCATCCTGGTTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCTTCCT 5340
Db  4989 CTGTTCCATCCTGGTTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCTTCCT 5048
Qy  5341 GCCTGTTCCTCCTCTGTGCTGCCCACTGTGGTGCCCACTGTGGTGCTCGTGGTCCAC 5400
Db  5049 GCCTGTTCCTCCTCTGTGCTGCCCACTGTGGTGCCCACTGTGGTGCTCGTGGTCCAC 5108
Qy  5401 ATTCTTCAGGATTCTCCGGAAGTCAGGCTTCCCGGCCCTCTCTTCTCTGCCCAGCAGCAGA 5460
Db  5109 ATTCTTCAGGATTCTCCGGAAGTCAGGCTTCCCGGCCCTCTCTTCTCTGCCCAGCAGCAGA 5168
Qy  5461 TCACGATTCCCGGAAGTCAGGCTTCCCGGCCCTCTCTTCTCTCTGCCCGGCA 5520
Db  5169 TCACGATTCCCGGAAGTCAGGCTTCCCGGCCCTCTCTTCTCTCTGCCCGGCA 5228
Qy  5521 CTCTTAGCAAACCTCCAGCAGGACTCCAGCACCCTACCCCACATAGACCTCTGACAGAGAAGCAGCACT 5580
Db  5229 CTCTTAGCAAACCTCCAGCAGGACTCCAGCACCCTACCCCACATAGACCTCTGACAGAGAAGCAGCACT 5288
Qy  5581 TTACATGGAGTCCTGGTCGGTGGGAGGCCATAGGCTACGGTGTAAAAGAGGCAGGGAAGTGGT 5640
Db  5289 TTACATGGAGTCCTGGTCGGTGGGAGGCCATAGGCTACGGTGTAAAAGAGGCAGGGAAGTGGT 5348
Qy  5641 GGTGTAGGAAAGTCAGGACTTCACATAGAGAGCCTAGCCCACACCAGAAATGACAGACAGA 5700
Db  5349 GGTGTAGGAAAGTCAGGACTTCACATAGAGAGCCTAGCCCACACCAGAAATGACAGACAGA 5408
Qy  5701 TCCCTCCTATCTCCCCCATAAGAGTTTGAGTCGAC 5735
Db  5409 TCCCTCCTATCTCCCCCATAAGAGTTTGAGTCGAC 5443
```

RESULT 3
ABL49724
ID   ABL49724 standard; DNA; 5443 BP.
XX
AC   ABL49724;
XX
DT   29-MAY-2002 (first entry)
XX
DE   Mouse promoter DNA sequence SEQ ID NO:2.
XX
KW   Human; p300; P300 transgenic animal; promoter; heart muscle cell;
KW   cardiac; heart failure; megalocardia; mouse; ds.
XX
OS   Mus musculus.
XX
PN   WO200205633-A1.
XX
PD   24-JAN-2002.
XX
PF   13-JUL-2001; 2001WO-JP006086.
XX
PR   14-JUL-2000; 2000JP-00215143.

```
Db  3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAGGGGGAAGGGAGAGCCTGGGATGGGAGC 3960
Qy  3961 TTGTGTGTGGAGGCAGGGGACAGACAGATATTGGAAGAGAAGGTGACCCTTACCCAG 4020
Db  3961 TTGTGTGTGGAGGCAGGGGACAGACAGATATTAAGCCTGGAAGAGAAGGTGACCCTTACCCAG 4020
Qy  4021 TTGTTCAACTCACCCTTCAGATTAAAATAACTGAGGTAAGGGCCTGGGTAGGGGAGGTG 4080
Db  4021 TTGTTCAACTCACCCTTCAGATTAAAATAACTGAGGTAAGGGCCTGGGTAGGGGAGGTG 4080
Qy  4081 GTGTGAGACGCTCCTGTCTCCTCCTGCATGCCCTGAGGCCCTTTGGGGAGGAGGAATGT 4140
Db  4081 GTGTGAGACGCTCCTGTCTCCTCCTGCATGCCCTGAGGCCCTTTGGGGAGGAGGAATGT 4140
Qy  4141 GCCCAAGGACTAAAAAAGGCCATGGCCAGAGGGGCCGAGGGCCAACGACTTTCATGG 4200
Db  4141 GCCCAAGGACTAAAAAAGGCCATGGCCAGAGGGGCCGAGGGCCAACGACTTTCATGG 4200
Qy  4201 GCAAACCTGGGGCCCGTAGTGATCGATGACAGAACTCGCCAATCGATACCCTTCTTC 4260
Db  4201 GCAAACCTGGGGCCCGTAGTGATCGATGACAGAACTCGCCAATCGATACCCTTCTTC 4260
Qy  4261 TTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCGAGTTTACCACTC 4320
Db  4261 GGAGCCAGGACAGGAGGGAAGTGGGAGGGAG 4291
Qy  4321 AAGTGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTC 4380
Db  4292 ------------------------------------------------------------ 4291
Qy  4381 CCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCGATGAT 4440
Db  4292 ------------------------------------------------------------ 4291
Qy  4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA 4500
Db  4292 ------------------------------------------------------------ 4291
Qy  4501 CCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGAT 4560
Db  4292 ------------------------------------------------------------ 4291
Qy  4561 AGAGAAAAGTGAAAGTCGAGTCTCGGTACCAGCAGGACTCCAAATTTAGGCAGCAGCA 4620
Db  4292 ---------------GGTCCCAGCAGGACTCCAAATTTAGGCAGCAGCA 4328
Qy  4621 TATGGGATGGGATATAAAGGGCTGGAGCACTGAGAGCTGTCAGAGATTTCTCCCAACCCA 4680
Db  4329 TATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGCTGTCAGAGATTTCTCCAACCCA 4388
Qy  4681 GGTAAGAGGGAGTTTCGGGTGGGGGTCTCTTCACCCACACCAGACCTCTCCCCACCTAGAA 4740
Db  4389 GGTAAGAGGGAGTTTCGGGTGGGGGTCTCTTCACCCACACCAGACCTCTCCCCACCTAGAA 4448
Qy  4741 GGAAACTGCCCTTTCCTGAAGTCAGGGTTCAGGGATCCTGACAGGGTGGCCTT 4800
Db  4449 GGAAACTGCCCTTTCCTGAAGTCAGGGTTCAGGGATCCTGACAGGGTGGCCTT 4508
Qy  4801 CCACCAGCCTGGGAAGTTCTCAGTGCAGCAGAGGTTTCCACAAGAAACACTGGAATGCCCCT 4860
Db  4509 CCACCAGCCTGGGAAGTTCTCAGTGCAGCAGAGGTTTCCACAAGAAACACTGGAATGCCCCT 4568
Qy  4861 TCCCTTACGCTGTCTTCCCATCTTCCTCCTGGATGCTCTCTGGGTCTTGGTTTATC 4920
Db  4569 TCCCTTACGCTGTCTTCCCATCTTCCTCCTGGATGCTCTCTGGGTCTTGGTTTATC 4628
Qy  4921 TTTGGCTCTTCGTCTTCGTCTTCAGCAAGATTGCCCTGTGCCTGTGCCTGTCCATCTTCCTCTACTGT 4980
Db  4629 TTTGGCTCTTCGTCTTCGTCTTCAGCAAGATTGCCCTGTGCCTGTGCCTGTCCATCTTCCTCTACTGT 4688
Qy  4981 CTCCGTGCCTGCCTGCCTTCTTGCGTGTCCTTCCTTCCACCCATTTCTCACTTCACC 5040
```

XX
PA   (CHUS) CHUGAI SEIYAKU KK.
PA   (HASE/) HASEGAWA K.
XX
PI   Hasegawa K, Kawase Y, Suzuki H;
XX
DR   WPI; 2002-179739/23.
XX
PT   Transgenic animals transferred with DNA encoding p300 and promoter
PT   exerting its activity in heart muscle cells, useful in studying onset
PT   mechanism of and screening remedies for heart failure.
XX
PS   Example 1; Page 47-53; 60pp; Japanese.
XX
CC   The present invention describes a transgenic animal transferred with a
CC   DNA encoding P300 and a promoter exerting its activity in heart muscle
CC   cells. Also described are: (1) a method for screening substances capable
CC   of treating heart failure comprising administering a test substance to
CC   the transgenic animal, and confirming inhibition of megalocardia or not
CC   in the animal; (2) substances thus screened; and (3) remedies for heart
CC   failure containing the screened substances as active ingredient. The
CC   transgenic animal can be used for studying the onset mechanism of heart
CC   failure. It can also be used for screening remedies for heart failure.
CC   The present sequence represents a mouse promoter which is used in an
CC   example from the present invention
XX
SQ   Sequence 5443 BP; 1352 A; 1399 C; 1492 G; 1200 T; 0 U; 0 Other;

Query Match   88.0%; Score 5045; DB 6; Length 5443;
Best Local Similarity   93.9%;
Matches 5383; Conservative   0; Mismatches   60; Indels 292; Gaps   1;

```
Qy      1 GGATCCTGCAAGGTCACACAAGGGTCTTCCACCCACCAGGTGCCTAGTCTCAATTTCAGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGATCCTGCAAGGTCACACAAGGGTCTTCCACCCACCAGGTGCCTAGTCTCAATTTCAGT 60

Qy     61 TTCCATGCCTTGTCTTCACAATGCTGGCCTCCCAGAGCTAATTTGGACTTTGTTTTTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCCATGCCTTGTCTTCACAATGCTGGCCTCCCAGAGCTAATTTGGACTTTGTTTTTAT 120

Qy    121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGCTACCCAGCTCTAAGGGTGCCCCGTGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTCAAAACGGGCCTGAATGAGGAGTAGATCTTGTGCTACCCAGCTCTAAGGGTGCCCCGTGA 180

Qy    181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTAGGTGGAAGCAGAATAAAGCAATTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTAGGTGGAAGCAGAATAAAGCAATTT 240

Qy    241 TCCTTAAAGCCAAAATCCTGCCTCTTCCTTTGTGACTCTCTAGACCTGTACCCTGTCAGTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCTTAAAGCCAAAATCCTGCCTCTTCCTTTGTGACTCTCTAGACCTGTACCCTGTCAGTT 300

Qy    301 AGGGTGGGGAGGTGGGGGTCGGGGAAGAGAAGAGGTGGGAAGTGGCAAAAGCCGATCCCTAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGGGTGGGGAGGTGGGGGTCGGGGAAGAGAAGAGGTGGGAAGTGGCAAAAGCCGATCCCTAG 360

Qy    361 GGGCCCTGGAAGTCGGAGCCTCTAGTATCCTAGTATCCTGATAATATGCAGAGCTGGGCCAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCCCTGTGAAGTCGGAGCCTCTAGTATCCTAGTATCCTGATAATATGCAGAGCTGGGCCAC 420

Qy    421 AAACATAGCAAGAAGTGATACCTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAACATAGCAAGAAGTGATAATCCTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGTT 480

Qy    481 GAAACAGGATTTAGAGAAGCCTCTGAACTCACCTGAACTCTGAAGCTCATCCACCAAGACA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAAACAGGATTTAGAGAAGCCTCTGAACTCACCTGAACTCTGAAGCTCATCCACCAAGACA 540

Qy    541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTACGCTGATATCCTAGCTGATAATATGCAGAGCTGGGCCAC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTACGCTGATATCCTAGCTGATAATATGCAGAGCTGGGCCAC 600

Qy    601 AGAAGTCCTGGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCGGCAAAGGTATGACCCCC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGAAGTCCTGGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCGGCAAAGGTATGACCCCC 660

Qy    661 TCAGCAGATGTAGTAATGTCCCTTAGATCCCCAGCAGTCTCTAAGAGACATG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCAGCAGATGTAGTAATGTCCCTTAGATCCCCAGCAGTCTCTAAGAGACATG 720

Qy    721 GGATGAGAGATGTAGTCATGTGCATTCAAACACAGCTATCCACAGTGTCCCTTGCCCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGATGAGAGATGTAGTCATGTGCATTCAAACACAGCTATCCACAGTGTCCCTTGCCCC 780

Qy    781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTCTTCCTCCCATCCTCCCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTCTTCCTCCCATCCTCCCAG 840

Qy    841 GACACACCCCTGGTCTGCAGTATTCATTTCTTCCTTCACGTCCCCTCTGTGACTTCCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACACACCCCTGGTCTGCAGTATTCATTTCTTCCTTCACGTCCCCTCTGTGACTTCCAT 900

Qy    901 TTGCAAGGCTTTTGACCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGCAAGGCTTTTGACCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 960

Qy    961 GCCATCTCAAGAGAAAGCAGACAACAGCAGGGGACCAGATTTGGAAGGATCAGGAACTAAA 1020
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCCATCTCAAGAGAAAGCAGACAACAGCAGGGGACCAGATTTGGAAGGATCAGGAACTAAA 1020

Qy   1021 TCACTGGCGCGGCCTGGGGTGAGAAAAAGAGTGAGTGCCTCCGCTCGGCCTAAGCCAAGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCACTGGCGCGGCCTGGGGTGAGAAAAAGAGTGAGTGCCTCCGCTCCGCCTAAGCCAAGC 1080

Qy   1081 TAGTCCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAGTCCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140

Qy   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTTTGAGTCTCCAACCCACCTACGAC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCAACCCACCTACAGAC 1200

Qy   1201 ACACTCGTGTGTGCGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACTGAGCTAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACACTCGTGTGTGCGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACTGAGCTAG 1260

Qy   1261 GCAACCAGACATGGCCGTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCAACCAGACATGGCCGTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320

Qy   1321 CAGGGGAGGAGAACAGAGAACAGAGAAATAAGGAAGAGGAAGGCCAGTCAATCAGA 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CAGGGGAGGAGAACAGAGAACAGAGAAATAAGGAAGAGGAAGGCCAGTCAATCAGA 1380

Qy   1381 TGCAGTCAGAAGAGATGGGAAGCCAACACAGCTTGAGCGAGGAGAAACAGAAAAGGGAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCAGTCAGAAGAGATGGGAAGCCAACACAGCTTGAGCGAGGAGAAACAGAAAAGGGAG 1440

Qy   1441 AGATTCTGGGACTAAGGAGGCCACAGAGAAGAAGAGGCCCCCCAAGTCTCCTCTTT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AGATTCTGGGACTAAGGAGGCCACAGAGAAGAAGAGGCCCCCCAAGTCTCCTCTTT 1500

Qy   1501 ATACCCTCATCCGTCTCCCAATTAAGCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATACCCTCATCCGTCTCCCAATTAAGCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560

Qy   1561 AGCGAAGAGACCCCTAGGGAGGATCACACTGGATGAAGGAGATGTGTGGAGAAGTCCAGG 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGCGAAGAGACCCCTAGGGAGGATCACACTGGATGAAGGAGATGTGTGGAGAAGTCCAGG 1620

Qy   1621 GCAACCTAAGAGCCAGAGCCTAAAAGCAGAAGAGATAAGGTGCTTCAAAGGTGGCCAGG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCAACCTAAGAGCCAGAGCCTAAAAGCAGAAGAGATAAGGTGCTTCAAAGGTGGCCAGG 1680

Qy   1681 CTGTGCCACACAGAGGGTCGAGGACTGGTGTAGCCTCAAGATAAGGATGATGCTCAGA 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy  2821  TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGTGGGAGCTGTGGTAGCAGGGAGAGTGT  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGTGGGAGCTGTGGTAGCAGGGAGAGTGT  2880

Qy  2881  TCCTCGGGTGTGAGGGTGTAGGGGAAAGCCAGAGACAGGGGAGTCTCGGCTTTGTCTCCTGAA  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  TCCTGGGGTGTGAGGGTGTAGGGGAAAGCCAGAGACAGGGGAGTCTCGGCTTTGTCTCCTGAA  2940

Qy  2941  CACAATGTCTACTTAGTTATAACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  CACAATGTCTACTTAGTTATAACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC  3000

Qy  3001  TGAAAAGACAGCAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTGGGTGCTTGATGGTG  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TGAAAAGACAGCAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTGGGTGCTTGATGGTG  3060

Qy  3061  CCACAAGGAGGGCATGAGTGTGAGTATAAGGCCCAGGAGCGTTAGAGAGGGCACTTG  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  CCACAAGGAGGGCATGAGTGTGAGTATAAGGCCCAGGAGCGTTAGAGAGGGCACTTG  3120

Qy  3121  GGAAGGGGTCAGTCTGCAGAGCCCATCCATGGAATCTGGAGCCTGGGGCCAACTGGTG  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  GGAAGGGGTCAGTCTGCAGAGCCCATCCATGGAATCTGGAGCCTGGGGCCAACTGGTG  3180

Qy  3181  TAAATCTCTGGGCTTGCCAGGCCATTCAAAGCACCAGCAGCACTTTCTCCCTCAGCGGC  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  TAAATCTCTGGGCTTGCCAGGCCATTCAAAGCACCAGCAGCACTTTCTCCCTCAGCGGC  3240

Qy  3241  GGCGGAAGGGAGGCAACCCCCACTTATACCCTTTCTCCCCACCTCCCAGGATTAACACCT  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GGCGGAAGGGAGGCAACCCCCACTTATACCCTTTCTCCCCACCTCCCAGGATTAACACCT  3300

Qy  3301  CTGCCTTCCCCCTTCCCACCTCCCATCAGGAGTTGAGAGGTTGCAGAGGGAGGTAAAAA  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CTGCCTTCCCCCTTCCCACCTCCCATCAGGAGTTGCAGAGGGAGGTTGCAGAGGGAGGTAAAAA  3360

Qy  3361  CCTACAGTCCAAACATGGGGCATGGCCATGGCCAGATATATGTGACGAGTCAGGGCAAGAAA  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  CCTACAGTCCAAACATGTGCACGATATATGTGACGAGATATGGATCAGTATGTGTAGAGGCAAGAAA  3420

Qy  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTGTCTG  3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTGTCTG  3480

Qy  3481  ACTGAAAACGGGCCATGGCGTGTGCTGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCA  3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  ACTGAAAACGGGCCATGGCGTGTGCTGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCA  3540

Qy  3541  GGTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTAAAGAGAT  3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GGTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTAAAGAGAT  3600

Qy  3601  TGGATGATGCCTGCATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT  3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  TGGATGATGCCTGCATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT  3660

Qy  3661  CTAAGCTAAGGTAGCAAGTCTTTGGAGGACACCTGTCTAGAGATGTGGCAACAGAGAC  3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  CTAAGCTAAGGTAGCAAGTCTTTGGAGGACACCTGTCTAGAGATGTGGCAACAGAGAC  3720

Qy  3721  TACAGACAGTATCTGTACAGAGTAAGGAGGAGGAGGGGTGTAGAATTCTCTTACTA  3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  TACAGACAGTATCTGTACAGAGTAAGGAGGAGGAGGGGTGTAGAATTCTCTTACTA  3780

Qy  3781  TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGATGCTCCCTAGACATCATGACTT  3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGATGCTCCCTAGACATCATGACTT  3840

Qy  3841  TGTCTCTCGGGGAGCCAGCACTGTCTGGGAACTTCAGGTCTGTGGAACTTCAGGTGTCTGTGGAACTGGAGGCTCCCCTCA  3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTCTCTCGGGGAGCCAGCACTGTCTGTGGAACTTCAGGTCTGTGGAACTGGAGGCTCCCCTCA  3900
```

```
Db  1681  CTGTGCACACAGAGGGTCTGAGGACTGGTGGTAGAGCCTCAAGATAAGGAGGATGATGCTCAGA  1740

Qy  1741  ATGGGCGGGGGGGGGGATTCTCGGGGGGGGGGGGGGGATTCTCGGGGGGGGGGGGGGAGAGAGAGGCTGGAACAG  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  ATGGGCGGGGGGGGGGATTCTCGGGGGGGGGGGGGGGATTCTCGGGGGGGGGGGGGGAGAGAGAGGCTGGAACAG  1800

Qy  1801  AGAATCTGGAAGCGCTGGAAACGATACCATAAAGGGAAGAAACCCAGGCTACCTTTAGATG  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGAATCTGGAAGCGCTGGAAACGATACCATAAAGGGAAGAAACCCAGGCTACCTTTAGATG  1860

Qy  1861  TAAATCATGAAAGACAGGGAGAGGGAGCTGAGAGACTAGAAGGACCCCGGGGCAGA  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  TAAATCATGAAAGACAGGGAGAGGGAGCTGAGAGACTAGAAGGACCCCGGGGCAAGA  1920

Qy  1921  CATGGAAGCAAGGACAAGCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGAAGGCAGAG  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  CATGGAAGCAAGGACAAGCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGAAGGCAGAG  1980

Qy  1981  CCCTGGTATGAGCACCAGAACAGCAGGCTAGGGTTAATGTCGAGACAGGGAACAGAGAAG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  CCCTGGTATGAGCACCAGAACAGCAGGCTAGGGTTAATGTCGAGACAGGGAACAGAGAAG  2040

Qy  2041  GTAGACACAGGAACAGAGACGCGGGGGAGCGGGGGAGCGGGGGAGGCTCAGGTAACAAGGTGTCCTTCTCAC  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GTAGACACAGGAACAGAGACGCGGGGGAGCGGGGGAGGCTCAGGTAACAAGGTGTCCTTCTCAC  2100

Qy  2101  CTGTGGCCAGAGCGTCCATCTGTCCACATACTCTAGAACTCATCAGACTGCAGGGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  CTGTGGCCAGAGCGTCCATCTGTCCACATACTCTAGAACTCATCAGACTGCAGGGC  2160

Qy  2161  TGGCTTGGGAGGACAGCTGGAAAGATATGTGAGACAAGGGAGCAACAAGGGAGACAAGGGCCTAGGA  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TGGCTTGGGAGGACAGCTGGAAAGATATGTGAGAGCAAGGGAGCAACAAGGGAGACAAGGGCCTAGGA  2220

Qy  2221  AAGGAAGAGAGGGCAAACCAGGCCACACAACAAGCCAGCAGAGGGCCAGAACTGAGTTAACTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  AAGGAAGAGAGGGCAAACCAGGCCACACAACAAGCCAGCAGAGGGCCAGAACTGAGTTAACTC  2280

Qy  2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCATGA  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCATGA  2340

Qy  2341  GCCCCACTACCCATACCAAGTTGGCCTGAGTGGAGTTCCTAGGTTCCCTGAGGACAGAG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GCCCCACTACCCATACCAAGTTGGCCTGAGTGGAGTTCCTAGGTTCCCTGAGGACAGAG  2400

Qy  2401  CCTGGCCTTGTCTCTTGTCTTGGGACCCAAGCTGACCCAATGTTCTCAGTGTTCTTGAAT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CCTGGCCTTTGTCTCTTGGGACCCAAGCTGACCCAATGTTCTCAGTGTTCTTATCAT  2460

Qy  2461  GCCCTCAAGAGCTTGAGAACCAGGCAGTCGATGGCCATGGCAGTTAGGCATATTAGGCAGCT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  GCCCTCAAGAGCTTGAGAACCAGGCAGTCGATGGCCATGACATATTAGGCAGCT  2520

Qy  2521  TGCACAGGACCTCAAGTGACCTCAGAGACACCAGGGGTGGCCTTTATCC  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  TGCACAGGACCTCAAGTGACCTCAGAGACACCAGGGGTGGCCTTTATCC  2580

Qy  2581  CCAAAGAGCAACCATTTGGCTCAAGAGCAATGGGAAATGGGAATGCAAGGTTGAATCAGGTC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  CCAAAGAGCAACCATTTGGCTCAAGAGCAATGGGAAATGCAAGGTTGAATCAGGTC  2640

Qy  2641  CCTTCAAGAATACTGCATGCAAGACCTTAAGAGAGAGGGGTATGCTCCTGCCC  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  CCTTCAAGAATACTGCATGCAAGACCTTAAGAGAGAGGGGTATGCTCCTGCCC  2700

Qy  2701  CCACCCACCATAAGGGAGTGAACTATCCTCAGGGGGCTTGGGAGGACACCAC  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  CCACCCACCATAAGGGAGTGAACTATCCTCAGGGGGCTTGGGAGGACACCAC  2760

Qy  2761  ATTACTGAGAGTGCTGAGCCCCAGCCCGAGAAAAACTGACCGCCCGTGTCCTGCCCACTTCCTGCCCACCTCCACAC  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  ATTACTGAGAGTGCTGAGCCCCAGCCCGAGAAAAACTGACCGCCCGTGTCCTGCCCACTTCCTGCCCACCTCCACAC  2820
```

```
Qy  3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAGGGGAGAGGAGAGCCTGGGATGGGAGC 3960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAGGGGAGAGGAGAGCCTGGGATGGGAGC 3960

Qy  3961 TTGTGTGTTGGAGGCAGAGATATTAAGCCTGAAGAGAAGGTGACCCTTACCCAG 4020
         ||||||||||||||||||||||||||||||  |||||||||| || |||||||||
Db  3961 TTGTGTGTTGGAGGCAGAGATATTAAGCCTGGAGAGAGGTCACCCTTACCCAG 4020

Qy  4021 TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTTGGTAGGGAGGTG 4080
         |||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db  4021 TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTTGGTAGGGAGGTG 4080

Qy  4081 GTGTGAGACGCTCCGTCTCCTCGCATGCCCTGAGGCCCTTTGGGGAGGAGGAATGT 4140
         |||||||||||||||||||||| ||| |||||||  |||||||||||||||||||||
Db  4081 GTGTGAGACGCTCCGTCTCCTCCTATCTCGGCCCATCGGCCCTTTGGGGAGGAGGAATGT 4140

Qy  4141 GCCCAAGGACTAAAAAAAAGGCCATGGACCAGAGGGCAACAGACCTTTCATGG 4200
         ||||||||||||||||||| |||||||||||||||||||||||||||||||||
Db  4141 GCCCAAGGACTAAAAAAAAGGCCATGGACCAGAGGGCGAGGGCAACAGACCTTTCATGG 4200

Qy  4201 GCAAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAATCGATACCCTTCTTC 4260
         ||||||||||||||||| || |||    ||||||||||||||||  ||||
Db  4201 GCAAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCTAGATTGGTCTCCCCAGCCTCGTACTCCTCTTCCT 4260

Qy  4261 TTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGA 4320
         |||||||||||||||||||||||||||| | |
Db  4261 GGAGCCACAGGAGGGAACCATGGACAGGGCCAGAGGGCAACAGAGGAGAAAAGTGA 4320

Qy  4321 AAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAGTCGAGTTTACCACTC 4380
         |||||||||||||||||
Db  4261 GGAGCCACAGGAGGGAAGGAGGGAGGGAG 4291

Qy  4381 CCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAA 4440
         ||||||||||||||||
Db  4292 ------ 4292

Qy  4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAGAGTTTA 4500
         ||||||||||||||
Db  4292 ------ 4292

Qy  4501 CCACTCCCTATCAGTGATAGAGAAAAGTGAGTTTACCACTCCCTATCAGTGAT 4560
         ||||||||||||||||||
Db  4292 ------ 4292

Qy  4561 AGAGAAAAGTGAAAGTCGGGTACCAGCAGAGACTCCGAATTTAGGCAGCAGCAGCA 4620
         ||||||||||||||||
Db  4292 ------GGTCCCAGCAGAGACTCCAGATTTAGGCAGCAGCAGCA 4328

Qy  4621 TATGGGATGGGATATAAAGGGGGCTGGAGCTGAGCACTGACAGAGATTTCTCCAACCCA 4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4329 TATGGGATGGGATATAAAGGGGGCTGGGAGCTGAGCACTGACAGAGATTTCTCCAACCCA 4388

Qy  4681 GGTAAGGGGGAGTTTCGGGTGGGGGCTCTTCACCCACCACCAGACCTCTCCCCACCTAGAA 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4389 GGTAAGGGGGAGTTTCGGGTGGGGGCTCTTCACCCACCACCAGACCTCTCCCCACCTAGAA 4448

Qy  4741 GGAAACTTGCCTTTCCTGGAAGTGGGGTTGGGGTTCAGGCGCCGGTCGTCAGCAGGTGGCCTT 4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4449 GGAAACTGCCTTTCCTGGAAGTGGGGTTGGGGTTCAGGCGCCGGTCGTGACAGGGTGGCCTT 4508

Qy  4801 CCACCAGCCTGGGAAGTTCTCAGTGCGCAGGAGGTTTCCACAAGAAAACACTGGATGCCCCT 4860
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4509 CCACCAGCCTGGGAAGTTCTCAGTGCGGCAGGAGGTTTCCACAAGAAAACTGGATGCCCCT 4568

Qy  4861 TCCCTTACGGCTGTCTCTTCCCATCTTCCCATCTTCCCATCTTTGTTTATC 4920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4569 TCCCTTACGGCTGTCTTCGTCTTTCCTCCTCCCTGGGATGCTCCTCCCCGTCTTTGGTTTATC 4628

Qy  4921 TTTGGCTCTTCGTCTTCAGCAAGATTTGCCTTGCCCTGTGCTGTCTCCACTCCATCTTTCTCTACTGT 4980
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4629 TTTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGCTGCTCCCTGTGCTGCTCCTCCACTCCATCTTTCTCTACTGT 4688

Qy  4981 CTCCGTGCCTGCCTTGCCTTCCTTCCGTGTCCTTTCCACCCATTTCTCACTTCACC 5040
         ||||||||||||||||||||||||||||||||||||
Db  4689 CTCCGTGCCTGCCTTGCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCACC 4748
```

```
Qy  5041 TTTTCTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4749 TTTTCTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4808

Qy  5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4809 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4868

Qy  5161 CTGTGTGTCAGAGTGCTGAGAATCACACACCTGGGGTTCCCACCCCTTATGTGTAAACAATCTTCCA 5220
         ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  4869 CTGTGTGTCAGAGTGCTGAGAATCACACACCTGGGGTTCCCACCCCTTATGTGTAAACAATCTTCCA 4928

Qy  5221 GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTCTGCCTCGTGGTGCTCCAGCTGGTCCAC 5280
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4929 GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTGCGCCACTGGTGCTCCAGCTGGTCCAC 4988

Qy  5281 CTGTTCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCTTCCT 5340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4989 CTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCTTCCT 5048

Qy  5341 GCCTGTTCCTCCTCTGTCCAGCTGCGCCACTGGTGCTGGTGCTGGTGCTCCAGCTGGTCCAC 5400
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5049 GCCTGTTCCTCCTCTGTCCAGCTGCGCCACTGGTGCTGGTGCTGGTGCTCCAGCTGGTCCAC 5108

Qy  5401 ATTCTTCAGGATTCTCTGAAAGTAACCAGGTGAGAAATGTTTCCCCTGTAGACAGCAGA 5460
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5109 ATTCTTCAGGATTCTCTGAAAGTAACCAGGTGAGAAATGTTTCCCCTGTAGACAGCAGA 5168

Qy  5461 TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCCTTCTTCTGCCCAGCTGCCGGCA 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5169 TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCCTTCTCTTCTGCCCAGCTGCCCGGCA 5228

Qy  5521 CTCCTTAGCAAACCTCAGGCCCATAGGCTACGGTGAAAGGAGGCACT 5580
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  5229 CTCCTTAGCAAACCTCAGGCCCATAGGCTACGGTGAAAGGAGGCACT 5288

Qy  5581 TTACATGGAGTCCTGGTGGGAGAGCCATAGGCCTAGTGTAAAAGAGGCAGGGAAGTGGT 5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5289 TTACATGGAGTCCTGGTGGGAGAGCCATAGGCCTAGTGTAAAAGAGGCAGGGAAGTGGT 5348

Qy  5641 GGTGTAGAAAGTGAAAGTCGGGACTTCACATAGAAGCCTAGCCCACACCAGAAATGACAGACAGA 5700
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5349 GGTGTAGAAAGTGAAAGTCGGGACTTCACATAGAAGCCTAGCCCACACCAGAAATGACAGACAGA 5408

Qy  5701 TCCCCTCATCTCCCCCATAAGAGTTTGAGTCGAC 5735
         ||||||||||||||||||||||||||||||||||
Db  5409 TCCCCTCATCTCCCCCATAAGAGTTTGAGTCGAC 5443
```

RESULT 4
AAT31006
ID   AAT31006 standard; DNA; 1679 BP.
XX
AC   AAT31006;
XX
DT   26-SEP-1996 (first entry)
XX
DE   Mouse cardiac alpha myosin heavy chain promoter.
XX
KW   Gene therapy; hypoxia related enhancer element; HREE; ischaemia;
KW   reperfusion; promoter; alpha myosin heavy chain; alpha-MHC; ds.
XX
OS   Mus sp.
XX
PN   WO9620276-A1.
XX
PD   04-JUL-1996.
XX
PF   13-NOV-1995;   95WO-IB000996.
XX

```
Query Match             22.3%;  Score 1279.8;  DB 2;  Length 1679;
Best Local Similarity   82.2%;  Pred. No. 0;
Matches 1616;  Conservative    0;  Mismatches   57;  Indels  292;  Gaps   1;

Qy  3767 GAATTCTCTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCT 3826
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GAATTCTCTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCT 60

Qy  3827 AGACATCATGACTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTCGAGAGTA 3886
         |||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db    61 AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTCGAGAGTA 120

Qy  3887 GGAGGCTCCCCTCAGCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGGAGGGGAGAG 3946
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGAGGCTCCCCTCAGCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGGAGGGGAGAG 180

Qy  3947 CCTGGGGATGGGGAGCTTGTGT TGGAGGCAGGGGACAGATATTAAGCTCTGGAGAGAGG 4006
         |||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db   181 CCTGGGGATGGGGAGCTTGTGTGTGGAGGCAGGGGACAGATATTAAGCTCTGGAGAGAGG 240

Qy  4007 TGACCCTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCT 4066
         ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCT 300

Qy  4067 GGGTAGGGGAGCTTGGGTGTG      AGACGCTCCTCCTGTCTCTCCTATCGGCCCCTTTG 4126
         |||||||||||||||||||||      |||||||||||||||||||||||||||||||||
Db   301 GGGTAGGGGAGCTTGGGTGTGGGTGTGAGACGCTCCTCCTGTCTCTCCTATCGGCCCCTTTG 360

Qy  4127 TGACCCTACCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCT 4186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGGAGGAGGAGGAATGCCCAAGGACTAAAAAAAGGCCATGGAGCCAGAGGGGCGAGGGCAA 420

Qy  4187 CAGACCTTCATGGGCAAACCTTGGGCGCCCGTAGTGATCGATTGACAAGAACTCGCCAAT 4246
         ||||||||||||||||||||||||||
Db   421 CAGACCTTTCATGGGCAAACCTTGGGCGCCCGTAGTGATCGATTGCCATCCCAGCAGCCAAG 480

Qy  4247 CGATACCCCTTCTTCTTAACGGACAGGAGCCAGGACAGAGTGGGAGGAG----------- 4306
         ||||||||||||||||| ||||||||||||||||||||||||||||||
Db   481 GGATCAAAGGAGGAGGAGCCAGGACAGGAGCCAGGACAGAGTGGGAGGAGGGGAGGGGAG 525

Qy  4307 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4366
Db   526 ----------------------------------------------------------- 525

Qy  4367 CGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTA 4426
Db   526 ----------------------------------------------------------- 525

Qy  4427 TCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAGT 4486
         ||||||||||||||||||||||||
Db  4486 CCAGCTGCCCGGCCACTCTTAGCAAACCTTAGCAAACCTTAGCAAACCTTAGCAAACCTTGACA 1508

Qy  4487 GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4546
Db   526 ------------------------------------------------------------

Qy  4547 TCCCTATCAGTGATGATAGAGAAAGTGAAAGTCGAGTCGGTTACCAGCAGAGACTCCAAAT 4606
         ||||||||||||||||||||||||||||||||||||------GGTCCCAGCAGAGGACTCCAAAT
Db   526 --------------------------------------GGTCCCAGCAGAGGACTCCAAAT 548

Qy  4607 TTAGGCAGCAGGCATATGGGATGGGGATATAAAGGGGCTGAGCACTGAGAGCTGTCAGAG 4666
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   549 TTAGGCAGCAGGCATATGGGATGGGGATATAAAGGGGCTGAGCACTGAGAGCTGTCAGAG 608

Qy  4667 ATTTCTCCAACCCAGGTAGAGGGAGTTTCGGGTGGGGGGCTTCTTCACCCACACCAGACCT 4726
         |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db   609 ATTTCTCCAACCCAGGTAGAGGGAGTTTCGGGTGGGGGCTTCTTCACCCACCACCAGACCT 668

Qy  4727 CTCCCCACCTAGAAGAGGAAACTGCCTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT 4786
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   669 CTCCCCACCTAGAAGAGGAAACTGCCTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT 728

Qy  4787 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCGGAGGTTTCCACAAGAAA 4846
         |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db   729 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCGGAGGTTTCCACAAGAAA 788

Qy  4847 CACTGGATGCCCCTTCCCTTACGCTGTCTTCCCATCTTCCTCGTGGATGCTCCTCCCC 4906
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   789 CACTGGATGCCCCTTCCCTTACGCTGTCTTCCCATCTTCCTCGTGGATGCTCCTCCCC 848

Qy  4907 CGTCTTGGTTTATCTTGGCTCTTCAGCAAGATTTGCCCGTGTCCACTCCA 4966
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   849 CGTCTTGGTTTATCTTGGCTCTTCAGCAAGATTTGCCCGTGTCCACTCCA 908

Qy  4967 TCTTTCTCTACTGTCTCCGTGCCTTGCCTTCCTTCCTTCCTTCCTTCCACCCA 5026
         ||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db   909 TCTTTCTCTACTGTCTCCGTGCCTTGCCTTCCTTCCTTCCTTCCTTCCACCCA 968

Qy  5027 TTTCTCACTTCACCTTTCTCCCCTTCCATTGTATTCATCCTTCCTTCCTTCCT 5086
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   969 TTTCTCACTTCACCTTTCTCCCCTTCCATTGTATTCATCCTTCCTTCCTTCCT 1028

Qy  5087 TCCTTCCTTCCTTCCTCCTTCCTCGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATG 5146
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1029 TCCTTCCTTCCTTCCTCCTTCCTCGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATG 1088

Qy  5147 TTTCTCACTTCACCTTTCTCTGTGTCAGAGTGCTCAGAGTGCTGAGAATTAAGCCTGGGG 5206
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1089 TTTCTCACTTCACCTTTCTCTGTGTCAGAGTGCTCAGAGTGCTGAGAATTAAGCCTGGGG 1148

Qy  5207 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCGTCGTGGGGTCTCTCTTACCTTCCTCA 5266
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1149 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCGTCGTGGGGTCTCTCTTACCTTCCTCA 1208

Qy  5267 CCCCTGGCTGTCCTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 5326
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1209 CCCCTGGCTGTCCTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 1268

Qy  5327 CTACTCCCTTCCTTGCCTGTCCTGTGTTCCTCCTGCGCGCCACTGGTGTGCCTCGTTC 5386
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1269 CTACTCCCTTCCTTGCCTGTCCTGTGTTCCTCCTGCGCGCCACTGGTGTGCCTCGTTC 1328

Qy  5387 CAGCTGTGGTCCACATTCTTCAGGATTCTCTGAAAAGTTAACCAGGTGAGAATGTTTCCC 5446
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1329 CAGCTGTGGTCCACATTCTTCAGGATTCTCTGAAAAGTTAACCAGGTGAGAATGTTTCCC 1388

Qy  5447 CTGTAGACAGCAGATCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTTTCTCTGC 5506
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1389 CTGTAGACAGCAGATCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTTTCTCTGC 1448

Qy  5507 CCAGCTGCCCGGCCACTCTTAGCAAACCTTAGCAAACCTTAGCAAACCTTACCCCACATGACA 5566
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1449 CCAGCTGCCCGGCCACTCTTAGCAAACCTTAGCAAACCTTAGCAAACCTTACCCCACATGACA 1508
```

```
Qy  3767 GAATTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCT     3826
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1    GAATTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCT     60

Qy  3827 AGACATCATGACTTTGTCTCTGGGAGCCAGCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA  3886
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61   AGACATCATGACTTTGTCTCTGGGAGCCAGCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA  120

Qy  3887 GGAGGCTCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAAGGGAGAG      3946
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GGGAGGCTCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAAGGGAGAG     180

Qy  3947 CCTGGGATGGGAGCTTGTGTTGGAGGCAGGGGACAGATATTAAGCCTGGAAGAGAAGG     4006
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  CCTGGGATGGGAGCTTGTGTGTGTGTGGAGGCAGGGGACAGATATTAAGCCTGGAAGAGAAGG  240

Qy  4007 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAGGGGCCT    4066
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241  TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAGGGGCCT    300

Qy  4067 GGGTAGGGGAGGTGGTGTGAGACGTCTCCTGTCTCTCCTGCATGCCCTGAGGCCCTTTG    4126
         |||||||||||||||||||||||||||||||||||||||||
Db  301  GGGTAGGGGAGGTGGTGTGAGACGTCTCCTGTCTCTCCTATCTGCCCATCGGCCCTTTG    360

Qy  4127 GGGAGGAGGAATGTGCCCAAGGACTAAAAAAGGCCATGGAGCCAGAGGGCGAGGGGCAA    4186
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  361  GGGGAGGAGGAATGTGCCCAAGGACTAAAAAAAGGCCATGGAGCCCAGAGGGGCGAGGGGCAA 420

Qy  4187 CAGACCTTTCATGGGGCAAACCTTGGGCGCCCGTAGTGATCGATTGACAAGAACTCGCCAAT  4246
         |||||||||||||||||||||||||||||||||
Db  421  CAGACCTTTCATGGGGCAAACCTTGGGCGCCCGTGTGTCCTGTCTCCTGTCCATCCCTATCAGTG  480

Qy  4247 CGATACCCCTTCTTCTTCTACGGACAGGAGGAACTCGAGTTTACCACTCCTATCAGTG    4306
         |||
Db  481  GGATCAAAGGAGGAGCCAGGAGCAGGAGGGGAAGTGGGAGGGAG                   525

Qy  4307 ATAGAGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGT          4366
Db  526  -----                                                          525

Qy  4367 CGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTA    4426
Db  526  -----                                                          525

Qy  4427 TCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGT    4486
Db  526  -----                                                          525

Qy  4487 GAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC    4546
Db  526  -----                                                          525

Qy  4547 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGCTCGGTACCAGCAGGACTCCAAAT    4606
         |||||||||||||||||||
Db  526  ------GGTCCCAGCAGCAGGAGGACTCCAAAT                              548

Qy  4607 TTAGGCAGCAGGCATATGGGATGGATAAAGGGCTGAGACTCGTGTCAGAG            4666
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  549  TTAGGCAGCAGGCATATGGGATGGATAAAGGGCTGAGACTCGTGTCAGAG            608

Qy  4667 ATTTCTCCAACCCAGGTAAGAGGGGAGTTTCGGGTGGGGGGCTCTTCACCCACCAGACCT   4726
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  609  ATTTCTCCAACCCCAGGTAAGAGGGGAGTTTCGGGTGGGGGGCTCTTCACCCACCACCACCT  668

Qy  4727 CTCCCCACCTAGAAGGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT   4786
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  669  CTCCCCACCTAGAAGGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT   728

Qy  4787 GACAGGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAA   4846
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  729  GACAGGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAA   788
```

```
Qy  5567 GAGAAGCAGGCACTTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA   5626
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1509 GAGAAGCAGGCACTTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA   1568

Qy  5627 GGCAGGGAAGTGGTGGTGTAGGAAAGTCAGGAAAGTCAGGACTTCACATAGAGCCTAGCCCCACCAG 5686
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1569 GGCAGGGAAGTGGTGGTGTAGGAAAGTCAGGACTTCACATAGAGCCTAGCCCCACCAG    1628

Qy  5687 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT                   5731
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1629 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT                   1673
```

RESULT 5
ABV74357
ID   ABV74357 standard; DNA; 1679 BP.
XX
AC   ABV74357;
XX
DT   05-FEB-2003 (first entry)
XX
DE   Mouse cardiac alpha-myosin heavy chain promoter fragment.
XX
KW   Therapeutic delivery system; electrical pulse generator; brain;
KW   electrically responsive promoter; ischaemic injury; cardiac; kidney;
KW   arterial occlusive disease; coronary arterial disease; stroke; mouse;
KW   alpha myosin; promoter; ds.
XX
OS   Mus sp.
XX
PN   WO200249669-A2.
XX
PD   27-JUN-2002.
XX
PF   20-DEC-2001; 2001WO-US049845.
XX
PR   21-DEC-2000; 2000US-0257460P.
PR   20-AUG-2001; 2001US-0313926P.
XX
PA   (MEDT) MEDTRONIC INC.
XX
PI   Schu CA, Padua R, Bonner M, Donovan MG, Soykan O;
XX
DR   WPI; 2002-537680/57.
XX
PT   Therapeutic delivery system used for treating e.g. peripheral arterial
PT   occlusive disease or stroke comprises electrical pulse generator operably
PT   coupled with genetically engineered cells in mammalian tissue.
XX
PS   Disclosure; Page; 66pp + Sequence Listing; English.
XX
CC   The invention relates to a therapeutic delivery system comprising an
CC   electrical pulse generator operably coupled with genetically engineered
CC   cells in a mammalian tissue. The genetically engineered cells also
CC   comprise a target gene operably coupled to an electrically responsive
CC   promoter. The delivery system is useful for the production of
CC   therapeutically useful gene or protein products, in repairing tissue
CC   injury (e.g. ischaemic injury, damaged cardiac tissue, kidney tissue,
CC   brain tissue or endothelial tissue), in stimulating cells for controlled
CC   expression of therapeutically useful gene and protein sequences and for
CC   treating peripheral arterial occlusive disease, coronary arterial disease
CC   or stroke. The present sequence is that of the mouse cardiac alpha-myosin
CC   heavy chain promoter fragment, exemplary of cardiac specific promoter
CC   regions. Note: The sequence data for this patent is not represented in
CC   the printed specification but is based on sequence information supplied
CC   to Derwent by the European Patent Office
XX
SQ   Sequence 1679 BP; 342 A; 486 C; 418 G; 433 T; 0 U; 0 Other;

Query Match   22.3%; Score 1279.8; DB 6; Length 1679;
Best Local Similarity   82.2%; Pred. No. 0;
Matches 1616; Conservative   0; Mismatches   57; Indels 292; Gaps   1;

```
KW   electrical response enhancer element; pacemaker.
XX
OS   Mus sp.
XX
PN   US2003204206-A1.
XX
PD   30-OCT-2003.
XX
PF   20-DEC-2001; 2001US-00027655.
XX
PR   21-DEC-2000; 2000US-0257460P.
PR   20-AUG-2001; 2001US-0313926P.
XX
PA   (MEDT ) MEDTRONIC INC.
PA
XX
PI   Padua RA,  Schu CA,  Bonner MD,  Donovan MG,  Soykan O;
XX
DR   WPI; 2004-032680/03.
XX
PT   Therapeutic delivery system useful for regulating delivery of therapeutic
PT   proteins and nucleic acids, comprises electrical pulse generator coupled
PT   with genetically engineered cells in mammalian tissue.
XX
PS   Disclosure; SEQ ID NO 5; 39pp; English.
XX
CC   The invention relates to a therapeutic delivery system comprising an
CC   electrical pulse generator coupled with genetically engineered cells in
CC   mammalian tissue. The genetically engineered cells further include a
CC   target gene coupled to an expression vector comprising an electrical response
CC   also relates to an electrically responsive promoter. The invention
CC   enhancer element, a tissue specific promoter heterologous to the element
CC   and a coding sequence, an apparatus for testing cells comprising an upper
CC   plate electrode, a lower plate electrode and a porous membrane positioned
CC   between electrodes during operation, and a method of treating a patient
CC   comprising providing the patient with an electrical pulse generator
CC   coupled with genetically engineered cells in a patient tissue. The
CC   electrical pulse generator is a pacemaker. The method is used for
CC   regulating the delivery of therapeutic proteins and nucleic acids. The
CC   invention provides controlled and local delivery of therapeutically
CC   important gene or protein products. This sequence represents the mouse
CC   cardiac alpha-myosin heavy chain promoter region, used in the method of
CC   the invention.
XX
SQ   Sequence 1679 BP; 342 A; 486 C; 418 G; 433 T; 0 U; 0 Other;

Query Match             22.3%;  Score 1279.8;  DB 12;  Length 1679;
Best Local Similarity   82.2%;
Matches 1616;  Conservative    0;  Mismatches   57;  Indels  292;  Gaps     1;

Qy   3767 GAATTCTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGATGCTCTCCCT 3826
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAATTCTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGATGCTCTCCCT 60

Qy   3827 AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA 3886
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA 120

Qy   3887 GGAGGCTCCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGGAAGGGAGAG 3946
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGAGGCTCCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGGAAGGGAGAG 180

Qy   3947 CCTGGGATGGAGAGCTTGTGTTGGAGGGACAGATATTAAGCCTGGAAGAGAAGAAGG 4006
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGGGATGGAGAGCTTGTGTTGGAGGGACAGATATTAAGCCTGGAAGAGAAGAAGG 240

Qy   4007 TGACCCTTACCAGTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGTAAGGGCCT 4066
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGACCCTTACCAGTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGTAAGGGCCT 300

Qy   4067 GGGTAGGGGAGGTGGTGAGACCGTCTCCATGCCCTGAGCCCTTTG 4126
          |||||||||||||||||||||||||||||||||||||||
Db    301 GGGTAGGGGAGGTGGTGAGACCGTCTCCATCTGCTCTCTCCCTTTG 360
```

```
Qy   4847 CACTGGATGCCCCTTCCCTTAGGCTGTCTTCTCCATCTTCCTCTGGGGATGTCTCCCC 4906
          |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db    789 CACTGGATGCCCCTTCCCTTACGCTGTCTTCTCCATCTTCCTCTGGGGATGTCTCCCC 848

Qy   4907 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTGCCCGTGTCGTCCACTCCA 4966
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    849 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTGCCCGTGTCGTCCACTCCA 908

Qy   4967 TCTTTCTCTACTGTCTCCGTGCCTTGCCTTCTTGCGTGTCCTTTCCACCCA 5026
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    909 TCTTTCTCTACTGTCTCCGTGCCTTGCCTTCTTGCGTGTCCTTTCCACCCA 968

Qy   5027 TTTCTCACTTCACCCTTTCTCCTCCTTCTCATTGTATTCATCCTTCCTTCCTTCCTTCCT 5086
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    969 TTTCTCACTTCACCCTTTCTCCTCCTTCTCATTGTATTCATCCTTCCTTCCTTCCTTCCT 1028

Qy   5087 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 5146
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1029 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 1088

Qy   5147 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTTATG 5206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1089 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTTATG 1148

Qy   5207 TAAACAATCTTCAGTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTTACCTTCCTCA 5266
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1149 TAAACAATCTTCAGTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTTACCTTCCTCA 1208

Qy   5267 CCCCCTGGCTGTCCGTGTCCATCCTGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 5326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1209 CCCCCTGGCTGTCCGTGTCCATCCTGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 1268

Qy   5327 CTACTCCTCTTCCTGCCTCGTGTCCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTC 5386
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1269 CTACTCCTCTTCCTGCCTCGTGTCCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTC 1328

Qy   5387 CAGCTGTGTGGTCCACATTCTTCAGGATTCTTCAGGATTCTCTGAAAAGTTAACCAGGTGAGAATGTTTCCC 5446
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1329 CAGCTGTGTGGTCCACATTCTTCAGGATTCTTCAGGATTCTCTGAAAAGTTAACCAGGTGAGAATGTTTCCC 1388

Qy   5447 CTGTAGACAGCAGCAGAATCACGATTCTCCCGGAAGTCAGCGCTTCCCGGAAGCTTCCAGCCCTCCTTCTTCTCTGC 5506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1389 CTGTAGACAGCAGCAGAATCACGATTCTCCCGGAAGTCAGCGCTTCCCGGAAGCTTCCAGCCCTCCTTCTTCTCTGC 1448

Qy   5507 CCAGCTGCCGCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTCTGACA 5566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1449 CCAGCTGCCGCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTCTGACA 1508

Qy   5567 GAGAAGCAGGCACTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 5626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1509 GAGAAGCAGGCACTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 1568

Qy   5627 GGCAGGAGAAGTGGTGGTGTGAGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACACCAG 5686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1569 GGCAGGAGAAGTGGTGGTGTGAGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACACCAG 1628

Qy   5687 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 5731
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1629 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 1673

RESULT 6
ADL14241
ID   ADL14241 standard; DNA; 1679 BP.
XX
AC   ADL14241;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Mouse cardiac alpha-myosin heavy chain promoter region.
XX
KW   Mouse; cardiac alpha-myosin heavy chain promoter; ds;
KW   electrical pulse generator; electrically responsive promoter;
```

```
Qy  4127 GGGAGGAGGAATGTGCCCAAGGACTAAAAAAGGCCATGGAGCCAGGAGGGCCAGGGCAA 4186
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGGAGGAGGAATGTGCCCAAGGACTAAAAAAGGCCATGGAGCCAGGAGGGCCAGGGCAA  420

Qy  4187 CAGACCTTTCATGGGCAAACCTGGGGCCCGTAGTGATCGAATGACAAGAACTCGCCAAT 4246
         |||||||||||||||||||||||||||||
Db   421 CAGACCTTTCATGGGCAAACCTGGGGCCCCAAG                           480

Qy  4247 CGATACCCTTCTTCTTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTG 4306
                                   ||||
Db   481 GGATCAAAGGAGGAGGAGCCAGGACAGGAGGAAGTGGGAGGGAG------            525

Qy  4307 ATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT       4366
Db   526 ------                                                       526

Qy  4367 CGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAGTTTACCACTCCCTA       4426
Db   526 ------                                                       526

Qy  4427 TCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGT 4486
Db   526 ------                                                       526

Qy  4487 GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4546
Db   526 ------                                                       526

Qy  4547 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGCTCGGTACCAGCAGAGGACTCCAAAT 4606
                                                   ||||||||||||||||||||
Db   526 -------GGTCCCAGCAGAGGACTCCAAAT                                548

Qy  4607 TTAGGCAGCAGCCATATGGGATGGGATGGGATAATAAAGGGGCTGGAGCACTGAGAGCTGTCAGAG 4666
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   549 TTAGGCAGCAGCCATATGGGATGGGATGGGATAATAAAGGGGCTGGAGCACTGAGAGCTGTCAGAG  608

Qy  4667 ATTTTCTCCAACCCAGGTAAGAGGGAGTTTCGGGTGGGGGCTCTTCACCCACCAGGACCT 4726
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   609 ATTTTCTCCAACCCAGGTAAGAGGGAGTTTCGGGTGGGGGCTCTTCACCCACCAGGACCT  668

Qy  4727 CTCCCCACCTAGAAGAGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCGGTTCAGAGATCT 4786
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   669 CTCCCCACCTAGAAGAGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCGGTTCAGAGATCT  728

Qy  4787 GACAGAGGTGGCCTTCCACCAGCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAAA 4846
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   729 GACAGAGGTGGCCTTCCACCAGCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAAA  788

Qy  4847 CACTGGATGCCCCTTCCCTTCCCTTACGCTGTCTTCTCCATCTTCTTCCTCTGGGGATGCTCCTCCC 4906
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   789 CACTGGATGCCCCTTCCCTTCCCTTACGCTGTCTTCTCCATCTTCTTCCTCTGGGGATGCTCCTCCC  848

Qy  4907 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCGTCTTGCCTTGCCCTGGCTGCTGTGCTCCACTCCA 4966
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   849 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCGTCTTGCCTTGCCCTGGCTGCTGTGCTCCACTCCA  908

Qy  4967 TCTTTCTCTACTGTCTCCGTGCTTGCCTTGCCTTGCCTTCGTGTCCTTCTTCCTTTCCACCCA 5026
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   909 TCTTTCTCTACTGTCTCCGTGCTTGCCTTGCCTTGCCTTCGTGTCCTTCTTCCTTTCCACCCA  968

Qy  5027 TTTCTCACTTCACCTTTTCTCTCCCTTCCTTCCTTCTCATTGTATTCATCCTTCCTTCCTTCCTT 5086
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   969 TTTCTCACTTCACCTTTTCTCTCCCTTCCTTCCTTCTCATTGTATTCATCCTTCCTTCCTTCCT   1028

Qy  5087 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTT 5146
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1029 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT   1088

Qy  5147 TCCTTCCTTCCTGTGTCAGAGTGCTGAGAGTCGAGAGTGCTCAGAGTCACACCTCGGGTTCCCACCCTTATG 5206
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1089 TCCTTCCTTCCTGTGTCAGAGTGCTGAGAGTCGAGAGTGCTCAGAGTCACACCTCGGGTTCCCACCCTTATG  1148

Qy  5207 TAAACAATCTTCCAGTGAGCCAGCTTCAGTGCTGCTGGGTGCTCTCTTACCTTCCTCA 5266
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1149 TAAACAATCTTCCAGTGAGCCAGCTTCAGTGCTGCTGGGTGCTCTCTTACCTTCCTCA 1208

Qy  5267 CCCCTGGCTGTCCTGTTCCATCCTGGTCAGGATCTCAGTGCTGCTGGGTCAGGCCTCTG 5326
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1209 CCCCTGGCTGTCCTGTTCCATCCTGGTCAGGATCTCAGTGCTGCTGGGTCAGGCCTCTG 1268

Qy  5327 CTACTCCTTCCTTCCTGCCTGTTCCTCTCTGTCCAGCTGCGCCACTGTGCGTGCCTCGTTC 5386
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1269 CTACTCCTTCCTTCCTGCCTGTTCCTCTCTGTCCAGCTGCGCCACTGTGCGTGCCTCGTTC 1328

Qy  5387 CAGCTGTGTCGTCCACATTCTTCAGGATTCCTGAAAAGTAACCAGGTGAGAATGTTTCCC 5446
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1329 CAGCTGTGTCGTCCACATTCTTCAGGATTCCTGAAAAGTAACCAGGTGAGAATGTTTCCC 1388

Qy  5447 CTGTAGACAGCAGATCACGATTCTCCCGAAGCTTCAGCCTCCAGCCCTCTCTTCTCTGC 5506
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1389 CTGTAGACAGCAGATCACGATTCTCCCGAAGCTTCAGCCTCCAGCCCTCTCTTCTCTGC 1448

Qy  5507 CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACCATAGACCTCTGACA 5566
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1449 CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGCACCCCTTACCCCACCATAGACCTCTGACA 1508

Qy  5567 GAGAAGCAGGCACTTACATGGAGTCCTGGTGGAGAGCCATAGGCTACGGGTGTAAAAGA 5626
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1509 GAGAAGCAGGCACTTACATGGAGTCCTGGTGGAGAGCCATAGGCTACGGGTGTAAAAGA 1568

Qy  5627 GGCAGGGAAGTGGTGGTGTAGGAAAGTCAGGACTTCACATGAGAGCCTTAGCCCACCACCAG 5686
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1569 GGCAGGGAAGTGGTGGTGTAGGAAAGTCAGGACTTCACATGAGAGCCTTAGCCCACCACCAG 1628

Qy  5687 AAATGACAGACAGATCCCTCCTATCTCCCCCATAAGAGTTTGAGT 5731
         |||||||||||||||||||||||||||||||||||||||||||||
Db  1629 AAATGACAGACAGATCCCTCCTATCTCCCCCATAAGAGTTTGAGT 1673
```

RESULT 7
ADS97705
ID   ADS97705 standard; DNA; 5190 BP.
XX
AC   ADS97705;
XX
DT   16-DEC-2004   (first entry)
XX
DE   Rabbit alpha-myosin heavy chain promoter.
XX
KW   Rabbit; alpha-myosin; promoter; transgenic; cardiomyopathy; cardiant; ds.
XX
OS   Oryctolagus cuniculus.
XX
PN   WO2004082370-A2.
XX
PD   30-SEP-2004.
XX
PF   12-MAR-2004; 2004WO-US007740.
XX
PR   13-MAR-2003; 2003US-0454947P.
PR   11-MAR-2004; 2004US-01122334.
XX
PA   (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
PI   Robbins J;
XX
DR   WPI; 2004-718695/70.
XX
PT   New animal cell that is stably transformed with an expression cassette
PT   comprising a promoter and a heterologous nucleotide sequence operably
PT   linked to the promoter, useful in identifying anti-cardiopathic
PT   compounds.
XX
PS   Claim 1; SEQ ID NO 1; 88pp; English.
XX
CC   The present sequence is that of the rabbit alpha-myosin heavy chain gene

CC   promoter, which is an atria-preferred promoter. The promoter can be used
CC   in composition of the invention for altering cardiac-preferred
CC   expression in transgenic animals. A claimed transgenic rabbit has in its
CC   genome a stably incorporated expression cassette comprising: a promoter
CC   having e.g. the present sequence or at least 90% identity to it or
CC   comprising at least 50 contiguous nucleotides of it capable of initiating
CC   transcription in an animal cell; and a heterologous nucleotide sequence
CC   that is operably linked to the promoter. In the present case, the
CC   promoter is capable of initiating atria-preferred transcription. Other
CC   transgenic animals, e.g. mouse, dog, goat, pig, cow, monkey, chimpanzee
CC   and sheep, may also be produced. The transgenic animal exhibits altered
CC   cardiac preferred expression of the heterologous nucleotide sequence and
CC   may exhibit an altered susceptibility to cardiopathy (or cardiomyopathy).
CC   Such animals can be used to identify anti-cardiopathic compounds.
XX
SQ   Sequence 5190 BP; 1095 A; 1525 C; 1642 G; 928 T; 0 U; 0 Other;

Query Match   7.1%; Score 407.6; DB 13; Length 5190;
Best Local Similarity   58.8%; Pred. No. 1.5e-104;
Matches 1265; Conservative   0; Mismatches 699; Indels 188; Gaps 25;

```
Qy   2855 GTGGGAGCTGGTAGCAGGGAGAGAGTGTTCCTGGGTGTGGGGGAAAGCCAGAGC 2914
          ||| ||||||| | ||||||||  |  | |||||||||| |||   |||| |||||
Db   2895 G-GTGGGAGGGGAGCTGAGGCAGTATTCTGGGTGTGGGGGTGGGGGGAAAGCCAGGGC 2953

Qy   2915 AGGGAGTCGGCTTGTCTTCTCCTGGAACACAATGTCTACTTAGTTATAAACAGGCATGACCT 2974
          ||||||||| ||| ||||||| |||||||||||||||||||||||    |  ||||| |||||
Db   2954 AGGAGAGTGTGGCTTTGGCTTCTGAACACAATGTCTACTTAGTCACAACAGGCACGGCCT 3013

Qy   2975 GCTAAAGACCCAACATCTACGACCTCTGAAA--AGACAGCAGCCTGGACGACGAGGGGT 3032
          |       ||||| |  | |||||||||||     | |||||||||| | ||||||||
Db   3014 GTTGAAGACCCGACCTACCGCCTCTGAGACGGGCAGGATGCAGGAGGAGGAGT 3073

Qy   3033 GTCTCTGAGCCTTGGGTGCTTGATGGTGCCACAAAGGAGGGCATGGGTGAGTATAAGG 3092
          | | |||||| ||||||||||||   ||| || ||| |       ||||| ||| |||
Db   3074 ACCTTGAGGGCTGGGTGCTTGAGGACCCCCCCCCCCCCAAAGACAGCAGCATGGGAGTGTGAGC 3133

Qy   3093 CCCCAGGAGCGTTAGAGAAGGGCACTTGGGAAGGGGTCAGTCTGCGAGAGCCCTATCCAT 3152
          | | |||| |  || || |||| || ||||||||||||||| || | |||||||| |||
Db   3134 TCAGGAGGGGATAGCAGGCTTCTATCTGGGAAGGGGGCAGACTACAGAGCCCCTCCCCC 3193

Qy   3153 GGAATCTGAGCCTGGGCCAACTGGTGTAAATCTTCTGGGCCTGCCAGGCATTCAAAGCA 3212
          | |  ||| | |||| | |   ||||| ||| |||| ||  |||||||| |||| | |
Db   3194 CAACCCTGCACCCCCTGCAGCCCTGTGGAACGCAGAGCCCAGCCCAATTGGTGTAAAT 3253

Qy   3213 GCACCTGCATCCTCTGCCAGCCTGGGAAGGGAGCAACCCCCCACTTATACCCT 3272
          | || | | ||||| | |||||||||      |||| |  |||||
Db   3254 CCCCAGGCCTGCTGTCTGGCGCCATCCTGTGGCGCCTTCCCCTTCCACCCTCTGGCAGCCTG 3313

Qy   3273 TTCTCCCTCAGGATTAACACCTCGGCGCCTTCCCCTTCCACCTCCCATCAGGA 3332
          |  |||| ||| |   |||||| ||||| |   ||||||  |||  |||||||
Db   3314 GAGAGGCGGAAGGGAGGCGGCGCCCAGACCG--GGCCGAGGATTAGT 3373

Qy   3333 GTGGAGGGTTGCAGAGGGATGGGTAAAAACCTACATGTCCAAACATCATGGTGCACGATAT 3392
          | ||||||| |||||||  |||||||||||| |   |||  | ||  || | ||| | |
Db   3374 GCAGAAGGTGGCAGCAGGAGGGGTAAAAGCCTTTGCGTGCA------------------GATGT 3418

Qy   3393 ATGGATCAGTATGTGTAGAGGCAAGAAAGAAAATCTGACGGTTAATGTGT 3452
          | |||| ||| | ||| ||||| || | |||   ||| |||  ||
Db   3419 GAGCATCGGCCTGCTGGAAAGCACGCAGGGGATCTCTGTAAGCTGAAGTGAAA-------- 3472

Qy   3453 AAAGTCTGTGCATGTGTGTTCTGACTGAAAACGGGCATGCGTGCAGCTGTTCA 3512
          |        |          | |                       | | | |
Db   3473 --------------------------------------TCTGCATGTGTCTGGTCA 3492

Qy   3513 GTTCTGTGCGTGAGGTTACCAGACTGCAGGTTTGTGTAAATTGCCCAAGGCAAAGTGG 3572
          ||| |||| ||| | | | |||| ||||| |||||  |   |||| || ||||||| ||
Db   3493 AGTCTGCACCGCAGGGTGTGACCAAGGTGTGCCGCCAGGTGTCGGAGGCAGCGTG 3552

Qy   3573 GTGAATCCCTTCCATGGTTTAAAGAGATTGGATGATGGGCCTGCATCTCAAGGACCATGGA 3632
          ||||| | ||||| |  |  |   ||||| |    | ||| || | ||| |||||||||||
Db   3553 GAGAATTCCTTCCACTGGGGTATGGGTGATGGGTTTCATGAGGGTGCATGGGGAGCACCGGG 3612

Qy   3633 AAATAGAATGGACACTCTATATGTCTCTAAGCTAAGGTGGCTGC---AAATG 3692
          | | |  | || | || | ||  ||||   |||||| |||| ||     ||
Db   3613 GAAGTGCTGGGGCTCCAGATGCCATTCCGAGTGCACCTGAGCATCTCAGACA 3672

Qy   3693 CCTGTCTAGAGATGTGGGCAACAGAGACTACAGATCTGTACAGGAAGGAGAGA 3752
          || |   |   |||||  | || | |  |||||||  ||  | | |||  || |
Db   3673 CCTGCC-------TGGGAATAGAAGCCCAGGACAGCCCAGTGTGGGGGGCAGG 3724

Qy   3753 GAGGAGGGGGTGTAGAATTCTTTACTATCAAAGGGAGGGAACTGAGTCGTCGTCAAAG 3812
          |||||| ||| ||      |         | | |||     | | ||| |||  ||||
Db   3725 ACAGAG-----------------AGGGAAACTGAGTC--TGGAGCTGCGTAG 3757

Qy   3813 TGGATGCTCCCTAGACATCATCGATGACTTTGTCTCTGGGAAGCACTGTGGAACTTCA 3872
          ||  | |||  | || ||| || |      |||| | |||| ||||| | |  |||| |
Db   3758 TAGACGCCTTGCTTCCAGCTCACAGG-----------------------TCCA 3785

Qy   3873 GGTCTGAGAGATTAGGAGGCTCCCTCAGCCT-GAAGCTATGCAGATAGCCAGGGTTGAA 3931
          ||| |||||||  || | || || | |||  | |||| | | | ||| || ||| ||||
Db   3786 GGTTCAGAAGTGTCGGGGAGACACTTGCAGCCTGGGTCGTGGGGAA 3845

Qy   3932 AGGGGGAGGGGA-GAGCCTGGGATGGGGGAGCTTGTGTGTTGGAGGGACAGATATTA 3990
          |||||| ||||   ||||| ||||  || || | | | | ||||||||||| || |
Db   3932 ...
```

```
Qy   2099 ACCTGTGCCAGAGCCGTCCAT-CTGTGTCCACATACTTAGAGATGTTCATCAGACTGCAG 2157
          |||| ||||||||||| ||||  | || || | |    ||| |  ||| | |||| ||||
Db   2140 ACCTTTGCCAGAGTATCCGTCCGTGCCCTGCACACTCCAGCAGGGCCTTCCAGCAGAATGCGG 2199

Qy   2158 GGCTGGCTTGGGAGGCAGCTG---------GAAAGAGTATGTGAGAGCCAGGGG 2201
          |  ||| || |||| || |||         ||||  ||| ||||||  |||||
Db   2200 AGCTGATTTGGGTGGCAGTGGAGCAGCGTGGACAGCAGAGAGTGTATGGGAACGCCAGGA 2259

Qy   2202 GAGACAAGGAGGGGGCCTAGGAAAGAAGAGGGCAAACCAGGGCCACACAAGAGGGCAGAG 2261
          || |  |||| || |  ||| ||| | |||  ||  ||||  ||| ||| |||||||||
Db   2260 CAGGAGGGGGGGCCCAGGGAAG--GAGAGGAGCAAACCAGGGCCGAGGATTAGT 2314

Qy   2262 CCCAGAACTGAGTTAACTCCTTGTTGCATCTTCCTCATACCATACCAAGTTTGGCCTGCATTCT 2321
          || |||| ||||||   |||| | | |||  || |  | |||||| |||||||||| |||| |
Db   2315 CCCAGAGCCAGGCTAACCCCCTTCCTTCTACCTCTCCTAGGAGG--GCAGGCACTGCC 2372

Qy   2322 TGACCACATCCCCCATGAGCCCCAGCCCCCGCCCCTCCTCCACCTTCCCCAAGTTTGGCCTGCATTCT 2381
          |||||| |||||||  || | | ||| |||  | || ||||||||  |||  |||||||||||| |||
Db   2373 TGCCCACCCCCTCAGCCCCCCGCCCCCTCCTCTCCACCTCCCCAAGTTTGGTCTGCTCAGGGGCGCA 2432

Qy   2382 AGGTTC-CCTGAGGACCAGAGCCTGGCCTTTGTCTTTGGACCTGACCCAAGCTGACCCAA 2440
          ||||||  |||  | ||| ||| ||||||  | |  | |||||||||||  ||| |||||
Db   2433 GGGTTCTCCCGAGGACCAAAGCGTCGGTCTTTGTCACCTGACCTACCCAGGTCGACCTGG 2492

Qy   2441 TGTTCTCAGTAGCCTTGAATGCCCTCAAGAGCTTGAGAACCAGGCAGTGACATATTAGGC 2500
          | ||| ||| | ||||| ||||| |||||||||||  | | ||||  || | | | || |
Db   2493 TATTC--CAGCCCCTTATCAGGTCCCAGGGGATGGGGCAACCAGGAGGGCACCAGGC 2551

Qy   2501 CATGGGCTAACCTGAGCTTGCACACAGGAGGAGCCTCAGGTGACCTCCAGGGACACAGCTG 2560
          |||||| |||||| | ||  | || ||  |   || |||| | ||| |||  |||| | |
Db   2552 CACGGGGATAACCC--CCAGCTTGTGCACTGGAGGCCTCAAGGTGACCT------AGAGCTA 2602

Qy   2561 CAGGACAGGTGGCCTTTATCCCCAAAGAGCAACCATTGGCATAGGTGGCTGC---AAATG 2617
          |||| |||||| ||  |  |||  |||| |||||||  || | ||||  ||      |||
Db   2603 GAGACAGGTGGCCTGCATCCCCGAGGGGCAACGCGGTTTGGCATGTGCACCTGAGGGATG 2662

Qy   2618 GGAATGCAAGGTTGAATCAGGTCCCTTCAAGAATACTGTCATGCAAAGACCTAAGACCCCTG 2677
          ||||| || |||||| | |||| | | || |||| ||  |  ||  || | |||||||||
Db   2663 GGAATGGCCGGTCAACTCAGGTCCCTCCAAGAGTCTCCACCCAGGAGAAGCTGACC 2722

Qy   2678 GAGAGAGGGGTATGCTCCTGCTCCCCAAAGAGCAACCATTAAGGGAGTGAACTATCCTAGGGGGGC 2737
          ||||| || ||| ||| ||   | || ||  |||||||||  |    | | |  || ||||||||
Db   2723 GAGAGAGATGTGTCTCTGCACTGGCGGCAACATCCTAGCGGGGC-----TAATAGGCAGCCAGCATCCTAGGGGGGC 2774

Qy   2738 TGGCGACCTTGGGGAGGAGACACCA--CATTACTGAGTGGCTGAG-CCCAGAAAAAACTGACC 2794
          |||| ||||  ||| ||||| || |   |  | || |||  ||||  |||    |  |||||
Db   2775 TGACAACCTTGGGGGAGGGAACCCATCTTCACGGACAGTGTCCCACCCAGAGAAGCTGACC 2834

Qy   2795 GCCCTGTGTCCCCACCTCCACACTCTAGAGCTCTAGTATTGAGAGGTGACAGTAGATAGG 2854
          || | | || | || |  || |||| |||  |   |  | |||||| |||  ||  | ||
Db   2835 CCTGCCTGTCCTCCTACCCCACTTCCAACTCTGGAGCTATATCGAGAGGTGTCAGTGGATAGG 2894
```

```
Db   3846  TGGGAGGAGGGTAGGGACAAGGGAGGGAAACTGCAGTGGGGGGGCAGGGGACAAGCATTC  3905
           |||||| ||||||||| ||| |||| |||||    |||  |||||  |||   ||| |||
Qy   3991  AGCCTGGAGAGAGAGGTGACCCTTACCCAGTGT--TCAACTCACCCTTCAGATTAAAAA   4048
           ||| |||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db   3906  GTCCTATATGAAAAGTGACCCTGCTCAACTCACCCTTCAGTTAAAA              3965

Qy   4049  TAACTGAGGTAAGGGCCTGGGTAGGGAGGTGTGT----GAGACGCTCCTGTCTCTCCT    4104
           |||||||||||||||||| |||||||||||||||    |||||||||||||||||||||
Db   3966  TAACCGAGGTAAAGGCCATGGTGGGCGGGGGAGGTGGTGAGAAGTCCTGTCTTCCCA   4025

Qy   4105  CTGCCATGCCCTGAGGCCCTTTGGGAGGAGGAATGTGCCCAAGGACTAAAAAAAGGCCAT  4164
           ||||   ||||| ||||||||  |||||||||||||||||   |||||||||||||||||
Db   4026  CTATCTGCCCATCAGCGCTCTGAGGGCCGAATGTGCTCAAGGACTAAAAAAAGGCCCT    4085

Qy   4165  GGAGCCAGAGGGGCGCAGGGCAACAGACCTTTCATGGGCAAACCTTGGGGCC        4216
           |||||  |||||| |||   |||||||||||||||||||   ||| |||||||
Db   4086  GGAGCCCGAGGGGCTGGGGCGCAGCAGCACCTTTCATGGGCAAATCTGGGGGCC      4137
```

RESULT 8

```
AAT11359
ID   AAT11359  standard; DNA; 520 BP.
XX
AC   AAT11359;
XX
DT   07-JUL-1996  (first entry)
XX
DE   Bidirectional promoter.
XX
KW   Tet repressor; Herpes simplex virus; virion protein 16; VP16;
KW   fusion protein; gene expression; regulation; inhibition; activation;
KW   transcription; ds.
XX
OS   Synthetic.
XX
PN   WO9601313-A1.
XX
PD   18-JAN-1996.
XX
PF   29-JUN-1995;  95WO-US008179.
XX
PR   01-JUL-1994;  94US-00270637.
PR   15-JUL-1994;  94US-00275876.
PR   03-FEB-1995;  95US-00383754.
PR   07-JUN-1995;  95US-00486814.
XX
PA   (BUJA/) BUJARD H.
PA   (GOSS/) GOSSEN M.
XX
PI   Bujard H, Gossen M;
XX
DR   WPI; 1996-087666/09.
XX
PT   New tetracycline-regulated transcription modulators - comprising fusion
PT   proteins which bind to tet operator sequences to activate or inhibit
PT   transcription.
XX
PS   Disclosure; Page 73; 112pp; English.
XX
CC   Fusion proteins comprising a first polypeptide which binds to a tet
CC   operator sequence in the presence of tetracycline or a tetracycline
CC   analogue, operatively linked to a second polypeptide which either
CC   activates or inhibits transcription in eukaryotic cells may be used to
CC   activate or inhibit transcription. Such proteins may be used to regulate
CC   gene expression in cells and may be particularly useful for gene therapy
CC   and for expression of gene products in transgenic organisms. Induction of
CC   gene expression is rapid, efficient and strong, typically 1000-2000 fold.
CC   The inducing agent does not cause pleitropic effects or cytotoxicity in
CC   eukaryotic cells. This sequence is a bidirectional promoter which can be
CC   used in the production of vector constructs. See also AAT11358
XX
SQ   Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;
```

Query Match   5.4%;   Score 311.8; DB 2;   Length 520;
Best Local Similarity   99.4%;   Pred. No. 8.7e-78;
    Matches 313;   Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
Qy   4277  GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC  4336
           ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63  GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC   122

Qy   4337  TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG  4396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123  TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG   182

Qy   4397  AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT  4456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183  AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT   242

Qy   4457  TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  4516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243  TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG   302

Qy   4517  ATAGAGAAAAGTGAAAGT  4576
           ||||||||||||||||||
Db    303  ATAGAGAAAAGTGAAAGT   362

Qy   4577  CGAGCTCGGTACCAG  4591
           |||||||||||||||
Db    363  CGAGCTCGGTACCCG   377
```

RESULT 9

```
AAT45722
ID   AAT45722  standard; DNA; 520 BP.
XX
AC   AAT45722;
XX
DT   18-MAR-1997  (first entry)
XX
DE   Bi-directional tetracycline-regulated promoter region.
XX
KW   Tet; tetracycline; Tc; operator; transcription; regulation; inducible;
KW   repressor; gene expression; therapy; transgenic animal; disease model;
KW   HSV; herpes simplex virus; tk; thymidine kinase; ds.
XX
OS   Synthetic.
XX
FH   Key               Location/Qualifiers
FT   promoter          24..427
FT                     /*tag= a
FT                     /note= "bi-directional, i.e. translation starts in a 3'
FT                     to 5' direction at posn. 24 and translation starts in 5'
FT                     to 3' direction at posn. 427, an explanatory figure is
FT                     given in the specification"
FT                     69..363
FT                     /*tag= b
FT                     /rpt_type= direct
FT                     /note= "seven repeats of the tet operator sequence
FT                     contained within the bi-directional promoter"
FT   repeat_region
XX
PN   WO9640892-A1.
XX
PD   19-DEC-1996.
XX
PF   06-JUN-1996;  96WO-US009049.
XX
PR   07-JUN-1995;  95US-00485971.
XX
PA   (BADI ) BASF AG.
XX
PI   Bujard H, Gossen M, Hillen W, Helbl V, Schnappinger D;
XX
DR   WPI; 1997-052305/05.
XX
```

```
PT   Nucleic acid encoding tetracycline-inducible transcription regulatory
PT   fusion protein - comprising modified tetracycline repressor able to bind
PT   mutant tet operator, fused to transcription regulator, useful for
PT   modulating eukaryotic gene expression.
XX
PS   Disclosure; Page 79-80; 117pp; English.
XX
CC   AAT45721 and AAT45722 are bi-directional, tetracycline (Tc)-regulated
CC   promoters used to co-ordinate regulation of expression of 2 genes of
CC   interest. The promoters contain tet operator sequences and a minimal
CC   herpes simplex virus (HSV) thymidine kinase (tk) promoter. The main
CC   invention of the specification concerns modified Tet repressor (TetR)
CC   proteins that bind to modified class B tet operator sequences teto-4C and
CC   teto-6C (see AAT45711 and AAT48478). Modified TetR proteins can be fused
CC   to any transcription regulatory polypeptide (e.g. HSV virion protein 16,
CC   VP16) and used to control transcription of a teto-4C or teto-6C linked
CC   gene. Nucleic acid encoding such a fusion protein may be introduced into
CC   a cell and transcription of the protein can be controlled by altering the
CC   concn. of tetracycline (or an analogue) in the cell, as appropriate. This
CC   ability to modulate gene expression in a predictable way is very useful
CC   in gene therapy and for recombinant protein prodn. in cultured cells or
CC   transgenic animals. The Tc-inducible system is also useful for the prodn.
CC   of transgenic' animal models for the study of disease and also for the
CC   study of gene function e.g during differentiation. The Tc-inducible
CC   system allows rapid activation of gene transcription without cellular
CC   toxicity, high concns. of inducer are not required
XX
SQ   Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match             5.4%;     Score 311.8;  DB 2;    Length 520;
Best Local Similarity  99.4%;     Pred. No. 8.7e-78;
Matches  313;   Conservative    0;   Mismatches    2;   Indels     0;   Gaps     0;

Qy    4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy    4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy    4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy    4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy    4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 4576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 362

Qy    4577 CGAGCTCGGTACCAG 4591
           |||||||||||||
Db     363 CGAGCTCGGTACCCG 377

RESULT 10
AAV60079
ID   AAV60079 standard; DNA; 520 BP.
XX
AC   AAV60079;
XX
DT   25-MAR-2003  (revised)
DT   04-DEC-1998  (first entry)
XX
DE   Bidirectional promoter region for coordinate regulation of 2 genes.
XX
KW   Tet repressor; tetracycline; regulation; expression;
KW   Tet operator-linked gene; bidirectional promoter region;
KW   coordinate regulation; tetracycline-regulated transcriptional activator;
KW   db.
```

```
XX
OS   Synthetic.
XX
PN   US5814618-A.
XX
PD   29-SEP-1998.
XX
PF   07-JUN-1995;    95US-00485978.
XX
PR   14-JUN-1993;    93US-00076327.
PR   14-JUN-1993;    93US-00076726.
PR   14-JUN-1994;    94US-00260452.
PR   01-JUL-1994;    94US-00270637.
PR   15-JUL-1994;    94US-00275876.
PR   03-FEB-1995;    95US-00383754.
XX
PA   (KNOL ) KNOLL AG.
PA   (BADI ) BASF AG.
XX
PI   Gossen M, Bujard H;
XX
DR   WPI; 1998-541795/46.
XX
PT   Tetracycline based regulation of gene expression - uses a tetracycline
PT   operator sequence joined to a gene of interest, the gene of interest
PT   being induced in the presence, but not absence of the antibiotic.
XX
PS   Disclosure; Fig 7B; 63pp; English.
XX
CC   The present sequence represents a bidirectional promoter region for
CC   coordinate regulation of 2 genes by a tetracycline-regulated
CC   transcriptional activator. The specification describes a method for
CC   regulating expression of a Tet (tetracycline) operator-linked gene in a
CC   cell of a subject. The method comprises introducing into the cell a
CC   nucleic acid encoding a fusion protein which inhibits transcription in
CC   eukaryotic cells, the fusion protein comprising a polypeptide which binds
CC   to a Tet operator sequence, operatively linked to heterologous second
CC   polypeptide which inhibits transcription in eukaryotic cells and
CC   modulating the concentration of a tetracycline (analogue) in the subject.
CC   The method is used for the regulation of gene expression system, using
CC   tetracycline (analogues). The system enables a gene coupled to the system
CC   to be induced in the presence of Tet and then stopped when Tet is
CC   removed. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match             5.4%;     Score 311.8;  DB 2;    Length 520;
Best Local Similarity  99.4%;     Pred. No. 8.7e-78;
Matches  313;   Conservative    0;   Mismatches    2;   Indels     0;   Gaps     0;

Qy    4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
           ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy    4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy    4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy    4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy    4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 4576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 362

Qy    4577 CGAGCTCGGTACCAG 4591
           |||||||||||||
Db     363 CGAGCTCGGTACCCG 377
```

```
Best Local Similarity  99.4%;  Pred. No. 8.7e-78;
Matches  313;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy     4277  GGAACTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCAC  4336
             |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63  GGATCTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCAC  122

Qy     4337  TCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG  4396
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      123  TCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG  182

Qy     4397  AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT  4456
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      183  AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT  242

Qy     4457  TACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  4516
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      243  TACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  302

Qy     4517  ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT  4576
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      303  ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT  362

Qy     4577  CGAGCTCGGTACCAG  4591
             |||||||||||||||
Db      363  CGAGCTCGGTACCCG  377
```

RESULT 11
AAX60045
ID   AAX60045 standard; DNA; 520 BP.
XX
AC   AAX60045;
XX
DT   20-MAR-2003 (revised)
DT   04-AUG-1999 (first entry)
XX
DE   Bidirectional promoter construct for regulation of 2 genes.
XX
KW   Transgenic mice; transgene; tet operator-linked gene; tetracycline;
KW   mouse-active transcriptional regulatory element; mutant Tet repressor;
KW   gene therapy; genetic disease; acquired disease; cancer; viral disease;
KW   vaccination; rheumatoid arthritis; hypopituitarism; wound healing;
KW   tissue regeneration; cancer; benign prostatic hypertrophy; hemophilia;
KW   erythrocytopenia; artherosclerosis; liver disease; Alzheimer's disease;
KW   Parkinson's disease; human disease model; ds.
XX
OS   Unidentified.
XX
PN   US5912411-A.
XX
PD   15-JUN-1999.
XX
PF   07-JUN-1995;   95US-00487472.
XX
PR   14-JUN-1993;   93US-00076327.
PR   14-JUN-1993;   93US-00076726.
PR   14-JUN-1994;   94US-00260452.
PR   01-JUL-1994;   94US-00270637.
PR   15-JUL-1994;   94US-00275876.
PR   03-FEB-1995;   95US-00383754.
XX
PA   (UYHE-) UNIV HEIDELBERG.
XX
PI   Bujard H, Gossen M;
XX
DR   WPI; 1999-357232/30.
XX
PT   Transgenic mice with inducible transgene activity useful for in vitro and
PT   in vivo protein production.
XX
PS   Disclosure; Fig 7B; 63pp; English.
XX
CC   The specification describes transgenic mice which have a transgene and a
CC   tet operator-linked gene integrated in the genome. The transgene
CC   comprises a mouse-active transcriptional regulatory element linked to a
CC   polynucleotide sequence that encodes a fusion protein which activates
CC   transcription of the tet operator-linked gene. The fusion protein
CC   comprises a mutated Tet repressor that binds a tet operator sequence in
CC   the presence of tetracycline (Tc) or a Tc analogue, linked to a
CC   polypeptide that activates transcription in eukaryotic cells. The
CC   transgenic system may be used for gene therapy to treat genes involved in
CC   genetic or acquired diseases. Gene therapy may be used to treat cancer,
CC   viral diseases, for vaccination, and to provide (Tc induced) regulated
CC   doses of a product (e.g. for the treatment or regulation of rheumatoid
CC   arthritis, hypopituitarism, wound healing and tissue regeneration,
CC   cancer, benign prostatic hypertrophy, hemophilia, erythrocytopenia,
CC   artherosclerosis and liver disease, Alzheimer's disease, and Parkinson's
CC   disease). The system may also be used to produce proteins in vivo (e.g.
CC   using mammalian, yeast or fungal cells) or in vitro (e.g. transgenic farm
CC   animals), to produce animal models of human disease, or to produce a
CC   stable cell line for gene cloning. The present sequence represents a
CC   bidirectional promoter construct used to control the regulation of two
CC   genes by a Tc-regulated transcriptional activator. (Updated on 20-MAR-
CC   2003 to correct PF field.)
XX
SQ   Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match   5.4%; Score 311.8; DB 2; Length 520;

RESULT 12
AAZ56126
ID   AAZ56126 standard; DNA; 520 BP.
XX
AC   AAZ56126;
XX
DT   27-MAR-2000 (first entry)
XX
DE   Bidirectional promoter region derived from CMV promoter.
XX
KW   Tetracycline; tetracycline repressor; gene expression regulation; cancer;
KW   gene therapy; arthritis; wound healing; tissue regeneration; promoter;
KW   ds.
XX
OS   Cytomegalovirus.
XX
PN   US6004941-A.
XX
PD   21-DEC-1999.
XX
PF   07-JUN-1995;   95US-00485740.
XX
PR   14-JUN-1993;   93US-00076327.
PR   14-JUN-1993;   93US-00076726.
PR   14-JUN-1994;   94US-00260452.
PR   01-JUL-1994;   94US-00270637.
PR   15-JUL-1994;   94US-00275876.
PR   03-FEB-1995;   95US-00383754.
XX
PA   (BADI) BASF AG.
PA   (BADI) BASF BIORESEARCH CORP.
PA   (KNOL) KNOLL AG.
XX
PI   Gossen M, Bujard H;
XX
DR   WPI; 2000-085798/07.
XX
PT   Regulation of gene expression in cells, useful for gene therapy of
PT   diseases, production of proteins in vitro and in vivo and production of
PT   stable cell lines for cloning.
XX
PS   Disclosure; Fig 7B; 64pp; English.
XX
CC   This sequence represents a bidirectional promoter, for use in the method
CC   of the invention. The invention relates to a method for regulation gene

CC   expression in a cell, using tetracycline-regulated fusion proteins. The
CC   method involves obtaining a cell from a subject, and introducing a
CC   nucleic acid molecule into the cell, which operatively links a gene to at
CC   least one tetracycline (tet) operator sequence. A second nucleic acid
CC   molecule is then introduced which encodes a fusion protein, where the
CC   fusion protein comprises a first polypeptide which binds to a tet
CC   operator sequence in the presence of tetracycline, or its analogue,
CC   operatively linked to a second polypeptide (e.g. VP16) which activates
CC   transcription in eukaryotic cells to form a modified cell. The modified
CC   cell can then be administered to the subject, and the concentration of
CC   tetracycline (or and analogue) can be regulated so that the expression of
CC   the gene is regulated. The method is useful for gene therapy of diseases
CC   such as cancer and arthritis or for tissue regeneration and wound
CC   healing. The method may also be useful for the production of proteins in
CC   vitro and in vivo and for the production of stable cell lines for cloning
XX
SQ   Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match   5.4%; Score 311.8; DB 3; Length 520;
Best Local Similarity   99.4%; Pred. No. 8.7e-78;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   63 GGATCTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 4576
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCAG 4591
           |||||||||||||
Db   363 CGAGCTCGGTACCCG 377

RESULT 13
AAH47631
ID   AAH47631 standard; DNA; 520 BP.
XX
AC   AAH47631;
XX
DT   30-NOV-2001 (first entry)
XX
DE   Nucleotide sequence of a bidirectional promoter region.
XX
KW   Fusion protein; tet operator; tetracycline; gene therapy; anti-cancer;
KW   rheumatoid arthritis; hypopituitarism; wound healing; hemophilia;
KW   diabetes; Alzheimer's disease; tet repressor; promoter; ds.
XX
OS   Unidentified.
XX
PN   US6271348-B1.
XX
PD   07-AUG-2001.
XX
PF   24-JAN-2000; 2000US-00489777.
XX
PR   14-JUN-1993;   93US-00076726.
PR   19-JUN-1993;   93US-00076327.
PR   14-JUN-1994;   94US-00260452.
PR   01-JUL-1994;   94US-00270637.

PR   15-JUL-1994;   94US-00275876.
PR   03-FEB-1995;   95US-00383754.
PR   07-JUN-1995;   95US-00485978.
PR   28-SEP-1998;   98US-00162184.
XX
PA   (BADI) BASF AG.
PA   (KNOL) KNOLL AG.
XX
PI   Bujard H, Gossen M;
XX
DR   WPI; 2001-556625/62.
XX
PT   Fusion protein for inhibiting transcription in eukaryotic cells useful in
PT   gene therapy applications comprises a first polypeptide, which binds to
PT   tet operator sequences, operatively linked to a heterologous second
PT   polypeptide.
XX
PS   Disclosure; Fig 7B; 69pp; English.
XX
CC   The invention relates to a fusion protein that comprises a first
CC   polypeptide which binds to tet operator sequences, operatively linked to
CC   a heterologous second polypeptide, which inhibits transcription in
CC   eukaryotic cells. The fusion proteins are tetracycline-responsive and are
CC   useful for regulation of transcription in eukaryotic cells and animals.
CC   The tetracycline (Tc)-controlled regulatory system is useful in various
CC   applications in gene therapy, such as in the treatment of various disease
CC   conditions e.g. rheumatoid arthritis, hypopituitarism, wound healing and
CC   tissue regeneration, anticancer treatments, benign prostatic hypertrophy,
CC   hemophilia, diabetes and artherosclerosis. They are also useful for bone
CC   marrow support therapy, treatment of central nervous system disorders
CC   e.g. Alzheimer's disease, Parkinson's disease (see AAH47628 for a
CC   detailed description of the uses). The present sequence represents the
CC   nucleotide sequence of a bidirectional promoter region for coordinate
CC   regulation of two genes of interest by tetracycline-regulated
CC   transcriptional activator
XX
SQ   Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match   5.4%; Score 311.8; DB 4; Length 520;
Best Local Similarity   99.4%; Pred. No. 8.7e-78;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   63 GGATCTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 4576
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCAG 4591
           |||||||||||||
Db   363 CGAGCTCGGTACCCG 377

RESULT 14
AAH25572
ID   AAH25572 standard; DNA; 520 BP.
XX
AC   AAH25572;
XX

```
DT  05-SEP-2001  (first entry)
XX
DE  Nucleotide sequence of a bidirectional promoter.
XX
KW  Transgenic plant; transgene; tet operator-linked gene; Tet repressor;
KW  tetracycline-regulated transcriptional regulatory system; tet operator;
KW  tetracycline; ds.
XX
OS  Unidentified.
XX
PN  US6242667-B1.
XX
PD  05-JUN-2001.
XX
PF  28-SEP-1998; 98US-00161902.
XX
PR  14-JUN-1993; 93US-00076327.
PR  14-JUN-1993; 93US-00076726.
PR  14-JUN-1994; 94US-00260452.
PR  01-JUL-1994; 94US-00270637.
PR  15-JUL-1994; 94US-00275876.
PR  03-FEB-1995; 95US-00383754.
PR  07-JUN-1995; 95US-00487472.
XX
PA  (BADI ) BASF AG.
PA  (KNOL ) KNOL AG.
XX
PI  Bujard H, Gossen M;
XX
DR  WPI; 2001-396837/42.
XX
PT  Transgenic plant for regulating the expression of genes, comprises a
PT  transgene integrated into the genome and a tet operator-linked gene in
PT  the genome.
XX
PS  Disclosure; Fig 7B; 65pp; English.
XX
CC  The specification describes a transgenic plant with a tetracycline-
CC  regulated transcriptional regulatory system. The transgenic plants have a
CC  transgene and a tet operator-linked gene integrated in the genome, which
CC  confers a detectable and functional phenotype on the plant when
CC  expressed. The transgene comprises a transcriptional regulatory element
CC  functional in cells of the plant operatively linked to a polynucleotide
CC  sequence encoding a fusion protein that activates transcription of the
CC  tet operator linked gene. The fusion protein comprises a first
CC  polypeptide that is a mutated Tet repressor that binds to a tet operator
CC  sequence in the presence of tetracycline or its analogue, operatively
CC  linked to a second polypeptide that activates transcription in eukaryotic
CC  cells. In the presence of tetracycline, the fusion protein binds to the
CC  tet operator-linked gene and activates transcription of the tet operator-
CC  linked gene such that it is expressed at a level sufficient to confer the
CC  detectable and functional phenotype on the plant. The level of expression
CC  of the tet-operator gene can be downmodulated by depleting tetracycline
CC  from the plant. The transgene can be used to regulate the expression of
CC  genes in the plant. The transgenic plant can be used to analyse the
CC  functions of cellular proteins. The present sequence represents a
CC  bidirectional promoter for coordinate regulation of two genes of interest
CC  by a tetracycline-regulated transcriptional activator. It is used to
CC  produce transgenic plants of the invention
XX
SQ  Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match             5.4%;  Score 311.8;  DB 5;  Length 520;
Best Local Similarity   99.4%;  Pred. No. 8.7e-78;
Matches  313;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy  4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTTTACCAC 4336
         ||  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTTTACCAC 122

Qy  4337 TCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182
```

```
Qy  4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTT 4456
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTT 242

Qy  4457 TACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy  4517 ATAGAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGT 4576
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAAGTGAAAGT 362

Qy  4577 CGAGCTCGGTACCAG 4591
         |||||||||||||||
Db   363 CGAGCTCGGTACCCG 377


RESULT 15
ACA94734
ID  ACA94734 standard; DNA; 520 BP.
XX
AC  ACA94734;
XX
DT  11-AUG-2003  (first entry)
XX
DE  Tet coordinated bi-directional promoter region #2.
XX
KW  Gene therapy; tet promoter; transgenic; rheumatoid arthritis; ds;
KW  hypopituitarism; wound healing; anti-cancer treatment; promoter;
KW  transgenic farm animal; stable cell line production; tetracycline.
XX
OS  Unidentified.
XX
PN  US2003022315-A1.
XX
PD  30-JAN-2003.
XX
PF  03-AUG-2001; 2001US-00921650.
XX
PR  14-JUN-1993; 93US-00076327.
PR  01-JUL-1994; 94US-00270637.
PR  15-JUL-1994; 94US-00275876.
PR  03-FEB-1995; 95US-00383754.
PR  07-JUN-1998; 95US-00485978.
PR  28-SEP-1998; 98US-00162184.
PR  24-JAN-2000; 2000US-00489777.
XX
PA  (BADI ) BASF AG.
PA  (KNOL ) KNOL AG.
XX
PI  Bujard H, Gossen M;
XX
DR  WPI; 2003-438975/41.
XX
PT  Regulating expression of a gene in cell of a subject, by utilizing
PT  components of tetracycline repressor/operator system of
PT  prokaryotes to regulate gene expression in eukaryotic cells.
XX
PS  Disclosure; Fig 7B; 71pp; English.
XX
CC  The invention relates to a method of regulating expression of a tet
CC  operator-linked gene in cell of a subject, which involves introducing
CC  into the cell a nucleic acid molecule encoding a fusion protein which
CC  inhibits transcription in eukaryotic cells, comprising a first
CC  polypeptide which binds to a tet operator sequence, operatively linked to
CC  a heterologous polypeptide which inhibits transcription in eukaryotic
CC  cells and modulating concentration of tetracycline, or its analogue in
CC  the subject. The method is useful for regulating expression of a
CC  exogenous or endogenous gene in a cell. The system has widespread
CC  applicability to the study of cellular development and differentiation in
CC  eukaryotic cells, plants and animals. For expression of e.g. oncogenes in
CC  eukaryotic cells in a controlled manner in cells to study their function.
```

CC   The system can be used to regulate the expression of site-specific
CC   recombinases such as CRE or FLP, to allow for irreversible modification
CC   of the genotype of a transgenic organism under controlled conditions at a
CC   particular stage of development. The method is also useful for gene
CC   therapy purposes, in treatment for either genetic or acquired diseases,
CC   including rheumatoid arthritis, hypopituitarism, wound healing and anti-
CC   cancer treatments, for large-scale production of proteins in vitro and in
CC   transgenic farm animals and for production of stable cell lines for gene
CC   cloning. The present sequence represents DNA encoding a tet promoter
XX
SQ   Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match   5.4%;   Score 311.8; DB 8;   Length 520;
Best Local Similarity 99.4%;   Pred. No. 8.7e-78;
Matches 313; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTTACCACTCCCTATCAGTGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTCCCTATCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCAG 4591
          |||||||||||||
Db   363 CGAGCTCGGTACCCG 377

Search completed: April 25, 2005, 03:43:34
Job time : 1993.22 secs

THIS PAGE BLANK (USPTO)

```
                GenCore version 5.1.6
    Copyright (c) 1993 - 2005 Compugen Ltd.
```

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 02:20:37 ; Search time 12059.9 Seconds
(without alignments)
18101.184 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatcctgcaaggtcacaca.........ccataagagtttgagtcgac 5735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Match	Query Length	DB	ID	Description	
c	1	534.4	9.3	613	8	AZ269661	AZ269661	RPCI-23-4
c	2	499.4	8.7	529	8	BH073204	BH073204	RPCI-24-3
c	3	473.8	8.3	496	8	AZ300332	AZ300332	RPCI-23-1
c	4	305.6	5.3	312	8	AZ457684	AZ457684	1M0261E12
	5	305.2	5.3	469	8	BH285252	BH285252	CH230-118
	6	266.4	4.6	447	8	AZ729312	AZ729312	RPCI-24-7
c	7	236	4.1	447	5	BY000251	BY000251	BY000251
	8	225.2	3.9	616	2	BF525096	BF525096	UI-R-AC0-
	9	145.4	2.5	370	8	AZ407663	AZ407663	1M0178C01
	10	133.4	2.3	390	5	BY000731	BY000731	BY000731
	11	133.4	2.3	409	5	BY061731	BY061731	BY061731
	12	133.4	2.3	415	5	BY083328	BY083328	BY083328
	13	133.4	2.3	739	6	CB596657	CB596657	AGENCOURT
	14	133.4	2.3	741	6	CB599099	CB599099	AGENCOURT
	15	133.4	2.3	783	6	CA490238	CA490238	AGENCOURT
	16	133.4	2.3	841	6	CA460851	CA460851	AGENCOURT
	17	133.4	2.3	843	6	CA495692	CA495692	AGENCOURT
	18	133.4	2.3	850	6	CA490440	CA490440	AGENCOURT
	19	133.4	2.3	917	6	CA490678	CA490678	AGENCOURT
	20	133.4	2.3	920	6	CA490677	CA490677	AGENCOURT
	21	132.8	2.3	871	6	CA490555	CA490555	AGENCOURT
	22	132.4	2.3	363	5	BY083663	BY083663	BY083663
	23	131.8	2.3	881	6	CB951696	CB951696	AGENCOURT
c	24	130	2.3	253	5	BX632021	BX632021	BX632021
c	25	130	2.3	459	1	AI642327	AI642327	vn86f10.y
	26	129.6	2.3	378	5	BY064144	BY064144	BY064144
	27	129.2	2.3	224	5	BY000659	BY000659	BY000659
c	28	129	2.2	707	1	AA855323	AA855323	vw71a12.r
	29	127	2.2	349	1	AI892571	AI892571	mr79f09.y
	30	126.8	2.2	412	5	BY435020	BY435020	BY435020
	31	126.8	2.2	468	5	BY414811	BY414811	BY414811
	32	126.2	2.2	243	5	BY000397	BY000397	BY000397
	33	125.2	2.2	368	2	BF472804	BF472804	UI-M-BZ0-
	34	125.2	2.2	385	2	BF422669	BF422669	UI-M-CG0p
c	35	125.2	2.2	428	2	BF462620	BF462620	UI-M-CG0p
	36	124.8	2.2	399	5	BY432934	BY432934	BY432934
	37	123.8	2.2	405	5	BY419314	BY419314	BY419314
	38	123.8	2.2	425	5	BY427154	BY427154	BY427154
	39	123.6	2.2	405	5	BY436012	BY436012	BY436012
c	40	123	2.1	426	1	AA140343	AA140343	mq89f06.r
c	41	123	2.1	473	1	AA919406	AA919406	vz13e04.r
	42	122.8	2.1	203	5	BY426544	BY426544	BY426544
	43	122.8	2.1	246	2	AV244652	AV244652	AV244652
c	44	122.8	2.1	278	4	BM124558	BM124558	L0541G01-
	45	122.8	2.1	313	5	BY415619	BY415619	BY415619

ALIGNMENTS

RESULT 1
AZ269661/c

```
LOCUS       AZ269661                 613 bp    DNA     linear   GSS 26-JUL-2000
DEFINITION  RPCI-23-444F20.TJ RPCI-23 Mus musculus genomic clone
            RPCI-23-444F20, genomic survey sequence.
ACCESSION   AZ269661
VERSION     AZ269661.1  GI:9483276
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 613)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-444F20.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 444  row: F  column: 20
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..613
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-444F20"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
                     EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
```

```
                        brain genomic DNA was isolated and partially digested
                        with a combination of EcoRI and EcoRI Methylase.  Size
                        selected DNA was cloned into the pBACe3.6 vector at the
                        EcoRI sites.  The ligation products were transformed into
                        DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

  Query Match         9.3%;  Score 534.4;  DB 8;  Length 613;
  Best Local Similarity 96.7%;  Pred. No. 6.2e-130;
  Matches 578;  Conservative 0;  Mismatches 16;  Indels 4;  Gaps 3;

Qy   3167 GGGCCAACTGGTGTAAATCTCTGGCCTGCCAGGCATTCAAAGC--AGCACCTGCATCC 3224
          ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||
Db    606 GGGCCAACCTGGTGTAAATCTCTGGGCCGCCCAGGCATTCAAAAGCAAGCACCTGCATCC 547

Qy   3225 TCTGGCCAGCCTGGGAGGCGGAGGGGAGCAACCCCCACTTATACCCTTTCTCCCCTCAGC 3284
          ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    546 TCTGGCCAGCCT-GGGAGGCGGAGGGAAGCAACCCCCACTTATACCCTTTCTCCCCTCAGC 488

Qy   3285 CCCAGGATTAAACACCTCTGGCCTTCCCCCTTCCACCTCCCATCAGGAGTGGAGGGTTGC 3344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    487 CCCAGGATTAAACACCTCTGGCCTTCCCCCTTCCACCTCCCATCAGGAGTGGAGGGTTGC 428

Qy   3345 AGAGGGAGGGTAAAAACCTACATGTCCAAACATCATGGTGCACGATATATGGATCAGTAT 3404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 AGAGGGAGGGTAAAAACCTACATGTCCAAACATCATGGTGCACGATATATGGATCAGTAT 368

Qy   3405 GTGTAGAGGCAAGAAAGGAAAATCGCAGGCTTAACTGGGTTAATGTGGTTAAAGTCTGTGTG 3464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 GTGTAGAGGCAAGAAAGGAAAATCGCAGGCTTAACTGGGTTAATGTGGTTAAAGTCTGTGTG 308

Qy   3465 CATGTGTGTGTGTCTGACTGAAAACGGGGCATGGCTGTGTGCAGTCTGTTCAGTTCTGTGCGTG 3524
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 CATGTGTGTGTGTCTGACTGAAAACGGGGCATGGCTGTGTGCAGTCTGTTCAGTTCTGTGCGTG 248

Qy   3525 AGGTTACCAGACTGCAGGTTTGTGTGTAAATGCCCAAGGCAAAGTG--GGTGAATCCCTT 3583
          ||||||||||||||||||||||||| ||||||||||||||||||||  ||||||||||||
Db    247 AGGTTACCAGACTGCAGGTTTGTGTGTAAATTGCCCAAGGCAAAGTGAGTGAAAAAATCCCTT 188

Qy   3584 CCATGGTTAAAGAGATTGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAATGG 3643
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 CCATGGTTAAAGAGATTGGATGATGGCCTGCATCTCAGGGACCATGGAAAATAGAATGG 128

Qy   3644 ACACTCTATATGTGTTCTCTCTAAGTCAAGGCTTTGGAGGACACCTGTCTAGAG 3703
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 ACACTCTATATGTGTCCCTCTAAGCCAAGGTAGCAAGGGTCTTTGGAGGACACCTGTCTAGAG 68

Qy   3704 ATGTGGGCAACAGAGACTACAGACAGTATCTGTACAGAGTAAGGAGAGAGAGGGGGG 3761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 ATGTGGGCAACAGAGACTACAGACAGTATCTGCACAGAGTAAGGAGAGAGAGAGGAGGGGGG 10

RESULT 2
BH073204/c
LOCUS       BH073204           529 bp    DNA      linear   GSS 18-JUL-2001
DEFINITION  RPCI-24-303B15.TV RPCI-24 Mus musculus genomic survey sequence.
            RPCI-24-303E15, genomic survey sequence.
ACCESSION   BH073204
VERSION     BH073204.1  GI:14892801
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
  COMMENT   Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
```

```
                        9712 Medical Center Dr., Rockville, MD 20850, USA
                        Tel: 301 838 0200
                        Fax: 301 838 0208
                        Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-24. For BAC
            library availability, please contact Pieter de Jong
            (pdejong@mail.cho.org). Clones may be purchased from BACPAC
            Resources (http://www.chori.org/bacpac/orderingframe.html). BAC end
            page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 303  row: E  column: 15
            Seq primer: T7
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source          1..529
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-24-303E15"
                     /sex="Male"
                     /cell_type="Spleen/Brain"
                     /clone_lib="RPCI-24"
                     /note=vector: pTARBAC1; Site 1: BamHI; Site_2: BamHI;
                     RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                     library was cloned in the pTARBAC1 cloning vector at the
                     BamHI sites using MboI partially digested male C57BL/6J
                     DNA."

ORIGIN

  Query Match         8.7%;  Score 499.4;  DB 8;  Length 529;
  Best Local Similarity 97.7%;  Pred. No. 1.2e-120;
  Matches 517;  Conservative 0;  Mismatches 11;  Indels 1;  Gaps 1;

Qy   2832 ATTGAGAGGTGACAGTAGATAGGGTGGGAGAGCTGGTAGCAGGGAGAGTGTTCCTGGGTGTG 2891
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 ATTGAGAGGTGCCAGTAGATAGGGTGGGAGAGCTGGTAGCAGGGAGAGAAAGTGTTCCTGGGTGTG 470

Qy   2892 AGGGTGTAGGGGAAAGCCAGAGAGCAGGGAGTCTGGCTTTGTCTCCTGACACAATGTCTA 2951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    469 AGGGTGTAGGGGAAAGCCAGAGAGCAGGGAGTCTGGCTTTGTCTCCTGACACAATGTCTA 410

Qy   2952 CTTAGTTATAACGGCATGACCTGCTAAAGACCCAACATCTACGACCTCTGAAAAGACAG 3011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 CTTAGTTATAACGGCATGACCTGCTAAAGACCCAACATCTACGACCTCTGAAAAGACAG 350

Qy   3012 CAGCCCTGGAGGACAGGGGTGTCTCGAGCGGGGTTGGGTGCTTGATGGTGCCACAAAGGAG 3071
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 CAGCCCTGGAGGACAGGGGTGTCTCGAGCGGGGTTGGGTGCTTGATGGTGCCACAAAGGAG 290

Qy   3072 GGCATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAAGGGCACTTGGGAAGGGGTCA 3131
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GGCATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAAGGGCACTTGGGAAGGGGTCA 230

Qy   3132 GTCTGCAGAGCCCCTATCCATGGAATCTGAGCCTGGGGGCCAACTGGTGTAAATCTCTGG 3191
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 GTCTGCAGAGCCCCTATCCATGGAATCTGAGCCTGGGGGCCAACTGGTGTAAATCTCTGG 170

Qy   3192 -GCCTGCCAGGCATTCAAAGCAGCACCTGCATCCTCTGGCCAGCCTGGGGAGGCGGAAGGG 3250
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 AAAAAGACAGGCATTCAAAGCAGCACCTGCATCCTCTGGCCAGCCTGGGGAGGCGGAAGGG 110

Qy   3251 AGCAACCCCCACTTATACCCTTTCTCCCCTCAGCCCCAGGATTAACACCTCTGGCCTTCC 3310
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 AGCAACCCCCACTTATACCCTTTCTCCCCTCAGCCCCAGGATTAACACCTCTGGCCTTCC 50

Qy   3311 CCCTTCCACCTCCCATCAGGAGTGGAGGGTTGCAGAGGGGTTGCAGAGGGAGGGTAAAA 3359
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 CCCTTCCACCTCCCATCAGGAGTGGAGGGTTGCAGAGGGTTGCAGAGGGAGGGGTAAAA 1

RESULT 3
AZ300332/c
LOCUS       AZ300332           496 bp    DNA      linear   GSS 27-JUL-2000
```

```
Db    257 TGTGCAGCTGTTCAGTTCTGTCGTGCGTGAGGTACCAGACTCGCAGGTTTGTGTGTAAATTGC 198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   3559 CCAAGGCAAAGTGGTGAATCCTTCCATGGTTTAAAGAGATTGGATGATGGCCTGCATC 3618
          |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    197 CCAAGGCAAAGTGGTGAACACCTTCCATGGTTTAAAGAGATTGGATGATGGCCTGCATC 138
Qy   3619 TCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCTCTAAGCTAAGGTAGCAAG 3678
          ||| |||||||||||||||||||||||||||||||||||||| |||||| ||||||||||
Db    137 TCAGGGACCATGGAAAATAGAATGGACACTCTATATGTGTCCTAAGCCAAGGTAGCAAG 78
Qy   3679 GTCTTTGGAGGACCACCTGTCTAGAGATGTGGCAACAGACTACAGACAGTATCTGTAC 3738
          ||||||||||||||||||||||||||||||||||||||||||||||||| |||| |||
Db     77 GTCTTTGGAGGACCACCTGTCTAGAGATGTGGCAACAGACTACAGACAGAATCTGCAC 18
Qy   3739 AGAGTAAGAGGAGAGAG 3755
          |||||||||||||||||
Db     17 AGAGTAAGAGGAGAGAG 1


RESULT 4
AZ457684/c
LOCUS       AZ457684                312 bp    DNA      linear    GSS 04-OCT-2000
DEFINITION  1M0261B12F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0261E12 F, genomic survey sequence.
ACCESSION   AZ457684
VERSION     AZ457684.1  GI:1615809
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 312)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D., Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000    Std Error: 0.00
            Plate: 0261  row: E   column: 12
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 312.
FEATURES             Location/Qualifiers
     source          1..312
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0261E12"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
```

```
DEFINITION  RPCI-23-117N22.TJ RPCI-23 Mus musculus genomic clone
            RPCI-23-117N22, genomic survey sequence.
ACCESSION   AZ300332
VERSION     AZ300332.1  GI:9542117
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 496)
  AUTHORS   Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S.,
            Akinret,B., Levins,M., Mcgann,S., Tsegaye,G., Geer,K., Krol,M., de
            Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-23
  JOURNAL   Unpublished (1999)
COMMENT     Other_GSSs: RPCI-23-117N22.TV
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the mouse BAC library RPCI-23. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
            or from Resea ch Genetics (info@resgen.com). BAC end page:
            http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
            Plate: 117  row: N  column: 22
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..496
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="RPCI-23-117N22"
                     /sex="Female"
                     /lab_host="DH10B"
                     /clone_lib="RPCI-23"
                     /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
                     EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
                     brain genomic DNA was isolated and partially digested
                     with a combination of EcoRI and EcoRI Methylase.  Size
                     selected DNA was cloned into the pBACe3.6 vector at the
                     EcoRI sites. The ligation products were transformed into
                     DH10B  electrocompetent cells (BRL Life Technologies)."
ORIGIN

Query Match       8.3%;  Score 473.8;  DB 8;  Length 496;
Best Local Similarity 98.4%;  Pred. No. 7.1e-114;
Matches  489;  Conservative  0;  Mismatches  7;  Indels  1;  Gaps  1;

Qy   3259 CCCACTTATACCCTTCTCCCTCAGCCCCAGGATTAACACCTCGGCCTTCCCCCTTCCC 3318
          |||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db    496 CCCACTTATACCCTTCTCCCTCAGCCCCAGGATTAACACCTCGGCCTTCCCCCTTCCC 437
Qy   3319 ACCTCCCATCAGGAGTGGAGGGTTGCAGAGGGAGGGTAAAAACCTACATGTCCAAACATC 3378
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    436 ACCTCCCATCAGGAGTGGAGGGTTGCAGAGGGAGGGTAAAAACCTACATGTCCAAACATC 377
Qy   3379 ATGTGCACGATATATGGATCAGTATGTGTAGAGGCAGGAAGAAATCTGCAGGCTTAA 3438
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    376 ATGTGCACGATATATGGATCAGTATGTGTAGAGGCAGGAAGAAATCTGCAGGC-TAA 318
Qy   3439 CTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTCTGACTGAAAACGGGCATGC 3498
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 CTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTCTGACTGAAAACGGGCATGC 258
Qy   3499 TGTGCAGCTGTTCAGTTCTGTCGTGCGTGAGGTTACCAGACTGCAGGTTTGTGTGTAAATTGC 3558
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
                electrophoresis. Vector DNA was prepared from a derivative
                of pWD42 (gi|4732114|gb|AF129072.1|, a copy-number
                inducible derivative of plasmid R1. The vector was ligated
                with adaptors complementary to the insert adaptors and
                purified. The sheared, adaptored mouse DNA was annealed to
                adaptored vector DNA, and transformed into
                chemically-competent E. coli XL10-Gold (Stratagene) cells
                and selected for ampicillin resistance."
ORIGIN

Query Match      5.3%;   Score 305.6; DB 8;  Length 312;
Best Local Similarity  98.7%;   Pred. No. 2.8e-69;
Matches 308;   Conservative   0;   Mismatches   4;   Indels   0;   Gaps   0;

Qy       605 GTCCTGGGGTGGTAGGAACTGACCAGTGACTTTCAGTGGCAAAGGTATGACCCCCTCAG 664
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       312 GTCCTGGGGTGGTAGGAACTGACCAGTGACTTTCAGTGGCAAAGGTATGACCCCCTCAG 253

Qy       665 CAGATGTAGTAATGTCCCCTTAGATCCCATCCCAGGCAGGTCTTCTAAGAGGACATGGGAT 724
             ||||||||||||||||||||||||||||||||||||||||||||| |||||||||| ||||
Db       252 CAGATGTAGTAATGTCCCCTTAGATCCCATCCCAGGCAGGTCTCTGAGAGGACACGGGAT 193

Qy       725 GAGAGATGAGTAGTCATGTGGCATTCCAAACACAGCTATCCACAGTGTCCTTGCCCCTTCC 784
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       192 GAGAGATGAGTAGTCAAGTGACATTCCAAACACAGCTATCCACAGTGTCCTTGCCCCTTCC 133

Qy       785 ACTTAGCCAGGAGGACGAGCAGTAAACCTTAGCCTATCTTCTTCCTTCCATCCTCCCAGGACA 844
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       132 ACTTAGCCAGGAGGACGAGCAGTAAACCTTAGCCTATCTTCTTCCTTCCATCCTCCCAGGACA 73

Qy       845 CACCCCCTGGTCTGCAGTATTCATTCTTCCTTCAGTCCCCTGTGACTTCCATTTGC 904
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        72 CACCCCCTGGTCTGCAGTATTCATTCTTCCTTCAGTCCCCTGTGACTTCCATTTGC 13

Qy       905 AAGGCTTTTGAC 916
             ||||||||||||
Db        12 AAGGCTTTTGAC 1
```

```
RESULT 5
BH285252
LOCUS       BH285252                     469 bp    DNA      linear    GSS 30-NOV-2001
DEFINITION  CH230-118C1.TJB CHORI-230 Segment 1 Rattus norvegicus genomic clone
            CH230-118C1, genomic survey sequence.
ACCESSION   BH285252
VERSION     BH285252.1  GI:17197658
KEYWORDS    GSS.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 469)
  AUTHORS   Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K.,
            Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D.,
            Riggs,F., de Jong,P. and Fraser,C.M.
  TITLE     Rat BAC End Sequences from Library CHORI-230 Segment 1 EcoRI segment
  JOURNAL   Unpublished (1999)
COMMENT     Other GSSs: CH230-118C1.TVB
            Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: szhao@tigr.org
            Clones are derived from the rat BAC library CHORI-230
            (http://www.chori.org/bacpac/rat230.html). For BAC library
            availability, please contact Pieter de Jong (pdejong@mail.cho.org).
            Clones may be purchased from BACPAC Resources
            (http://www.chori.org/bacpac/ or ering information.htm). BAC end
            page: http://www.tigr.org/tdb/bac_ends/rat/bac_ends/bac_end_intro.html
            Plate: 118 row: C column: 1
```

```
Seq primer: SP6
Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..469
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="BN/SsNHsd/MCW"
                     /db_xref="taxon:10116"
                     /clone="CH230-118C1"
                     /sex="Female"
                     /cell_type="Brain"
                     /clone_lib="CHORI-230 Segment 1"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
                     CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
                     Pieter de Jong"
ORIGIN

Query Match      5.3%;   Score 305.2; DB 8;  Length 469;
Best Local Similarity  89.2%;   Pred. No. 4.2e-69;
Matches 387;   Conservative   0;   Mismatches  38;   Indels   9;   Gaps   5;

Qy      3787 GGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCTAGACATCATGACTTTGTCTC 3846
             |||||||||||||||||||||||||||| ||||||||| |||||||||||||||||||||
Db         2 GGAAACTGAGTCATGCACCTGCAAAATGAATGCCTCCCTGGACATCATGACTTTGTCCC 61

Qy      3847 TGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTAGGAGGCTCCCCTCAGCCTGA 3906
             |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||
Db        62 TGGGGAGCCAGCACTGTGGAACTCCAGGTCT--GAGAGTAGGAGGCACCCCTCAGCCTGA 119

Qy      3907 AGCTATGCAGATAGCCAGGGTTGAAA--GGGGGAAGGAGGAGAGCCTGGGATGGGAGCTTGT 3964
             ||||||||||||||||||||||||||   ||| |||||||||||||||||||||||||||
Db       120 AGCTGTGCAGATAGCTAGGGTGTAAAAGAGGGGGAAGGAGGAGGGGTGGAATGGGATCTTGT 179

Qy      3965 GTGTTGGAGGCAGGGGACAGATATTAAGCCTGGAAGAAGAGGTGACCCTTACCCAGT-TG 4023
             ||||||||||||||||||||||||||||||||||| |||||||||||||||||||||| ||
Db       180 GTGTTGGAGACAGGGGACAAATATTAGGCCCGTAAGCAAGGTAAGGGCC---TGGGTAGGGGAGGGTG 239

Qy      4024 TTCAACTCACCCCTTCAGATTAAAAATAAAAAATAACTGAGTGAGGAGAGAGGGC--- 4080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       240 TTCAACTCAGCCTTCAGATTAAAATAAAAATAACTAAGGTAAGGGCCATGTGGGTAGGGGAGGGTG 299

Qy      4081 GTGTGAGACGCTCCGTCCTCCTCCCTCCGATGCCTGAGCCCTTTGGGAGGAGGAGGAATGT 4140
             |||||||||||| ||||| ||||||| |||||||| ||| |||
Db       300 GTGTGAGACGGTCCGTCTCTCCTCCTATCGTCCCATCGGCCCCTTTGGGAGGAGGAGGAATGT 359

Qy      4141 GCCCAAGGACTAAAAAAAGGCCATGAGGGCGAGGGGCAGAGGGCGAGGGCAACAGACCTTTCATGG 4200
             ||||||||||||||||||||| |||||||||| |||||||
Db       360 GCCCAAGGACTAAAAAAAGGCCCTGGAGCCAGAGGGCGAGGGGCAGCAGACCTTT-ATGG 418

Qy      4201 GCAAACCTTGGGGC 4214
             ||||||||||||||
Db       419 GCAAACCTTCAAGGC 432
```

```
RESULT 6
AZ729312
LOCUS       AZ729312                     447 bp    DNA      linear    GSS 25-JAN-2001
DEFINITION  RPCI-24-79L13.TVB RPCI-24 Mus musculus genomic clone RPCI-24-79L13,
            genomic survey sequence.
ACCESSION   AZ729312
VERSION     AZ729312.1  GI:12486941
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
            Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
            Russell,D., de Jong,P. and Fraser,C.M.
  TITLE     Mouse BAC End Sequences from Library RPCI-24
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shaying Zhao
```

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.chori.org/bacpac/orderingframe.htm) . BAC end
page: http://ww .tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 79 row: L column: 13
Seq primer: T7
Class: BAC ends.

FEATURES   Location/Qualifiers
     source   1..447
                  /organism="Mus musculus"
                  /mol_type="genomic DNA"
                  /strain="C57BL/6J"
                  /db_xref="taxon:10090"
                  /clone="RPCI-24-79L13"
                  /sex="Male"
                  /cell_type="Spleen/Brain"
                  /clone_lib="RPCI-24"
                  /note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;
                  RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
                  library was cloned in the pTARBAC1 cloning vector at the
                  BamH1 sites using MboI partially digested male C57BL/6J
                  DNA."

ORIGIN

Query Match   4.6%;   Score 266.4;   DB 8;   Length 447;
Best Local Similarity 97.6%;   Pred. No. 8.3e-59;
Matches   281; Conservative   0; Mismatches   6; Indels   1; Gaps   1;

Qy   1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCCACCAGGTGCCCTAGTCTCAATTTCAGT 60
          ||||||||||| ||| ||||||| ||||||||||||||||||||||||||||||||||||
Db   161 GGATCCTGCAAGG-CATACATGAGTCTCCACCCACCAGGTGCCCTAGTCTCAATTTCAGT 219

Qy   61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTTGGACTTTGTTTTTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   220 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTTGGACTTTGTTTTTAT 279

Qy   121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGCTACCCAGCTCTAAGGGTGCCCGTGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   280 TTCAAAAGGGCCTGAATGAGGAGTAGCTCTTGTGCTACCCAGCTCCAAGGGTGCCCGTGA 339

Qy   181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   340 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 399

Qy   241 TCCTTAAAGCCAAAATCCTGCCTCTAGACTCTTCTTCTTCTGACCTCGG 288
          |||||||||||||||||||||||||||||||||||||||||||||
Db   400 TCCTTAAAGCCAAAATCCTGCCTCTAGACTCTTCTTCTTCTGACCTTGG 447

RESULT 7
BY000251/c
LOCUS   BY000251   447 bp   mRNA   linear   EST 06-DEC-2002
DEFINITION BY000251 RIKEN full-length enriched, adult male heart Mus musculus
            cDNA clone 1020001I03 5', mRNA sequence.
ACCESSION   BY000251
VERSION   BY000251.1   GI:26060500
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 447)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,

Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
Kawaji,H., Kawasaw,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
FEATURES   Location/Qualifiers
     source   1..447
                  /organism="Mus musculus"
                  /mol_type="mRNA"
                  /db_xref="taxon:10090"
                  /clone="1020001I03"
                  /sex="male"
                  /tissue_type="heart"
                  /dev_stage="adult"
                  /lab_host="BNN132"
                  /note="Site 1: SalI; Site 2: BamHI; cDNA library was
                  prepared and sequenced in Mouse Genome Encyclopedia

FEATURES
Seq primer: M13 Forward.
　　　　　Location/Qualifiers
source　　1..616
　　　　　/organism="Rattus norvegicus"
　　　　　/mol_type="mRNA"
　　　　　/strain="Sprague-Dawley"
　　　　　/db_xref="taxon:10116"
　　　　　/clone="UI-R-AC0-yh-h-03-0-UI"
　　　　　/dev_stage="adult"
　　　　　/lab_host="DH10B (Life Technologies)"
　　　　　/clone_lib="UI-R-AC0"
　　　　　/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
　　　　　polylinker; Site_1: Not I, Site_2: Eco RI; The UI-R-AC0
　　　　　library is a non-normalized library constructed from 16.5
　　　　　dpc rat atrioventricular (AV) canal. The tag is a string
　　　　　of 5 nucleotides present between the Not I site and the
　　　　　oligo-dT track. The library was constructed as described
　　　　　by Bonaldo, Lennon and Soares, Genome Research 6:
　　　　　791-806, 1996. Tissue provided by Jim Lin, Department of
　　　　　Biology, University of Iowa."

ORIGIN

Query Match　　　3.9%;　　Score 225.2;　DB 2;　Length 616;
Best Local Similarity　73.0%;　Pred. No. 8.2e-48;
Matches　371;　Conservative　0;　Mismatches 128;　Indels　9;　Gaps　6;

Qy　5057　TTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTT　5116
　　　　　|||||| ||||| |||||||| ||| |||-|||-|||-|||||| ||||-||
Db　　　8　TCTGTCTCTACTCTCTCTGCCTTCTCTCTTGTGTGTTCTCCTTTTCACGTGTTTCTCA　67

Qy　5117　CTCCCTTCCTTCCTTCCTTCCTTCC--TTCCTTCCTTCCTTCCTGTGTCGAGAGTGC　5174
　　　　　| |||| |||||||| |||||||||||　||||||||||| |||-||||||||||
Db　　　68　CTCCACTGCCTCCCCCCCCCCCTTCATTTTATCTTTCATTTTCTGTCTGTCAGAAATGC　127

Qy　5175　TGAGAATCACACCTGGGGTTCCCACCC--TTATGTAAACAATCTTCAGTGAGCCACAGC　5232
　　　　　||||||||||||| ||| ||||| |||　|||| ||||| ||||||||||||||||
Db　　128　TGGGAATCAAACCCAGGGCTTCATACACGTCAAGTAAGCAATCTCCCAGTGAGCCAAAGC　187

Qy　5233　TTCAGTGCGTCGGGTGCTCCTTCACCCCTCACCCCCGGCTTGTCCGTTCCTCCATCCT　5292
　　　　　||||||-| ||||| ||||| |||||| ||||| ||||||| |||| |||| |||||
Db　188　TTTAATCCT-CTGGGTGCTGTCTTACCGACCTCACTCCCTGCTCTGTCTGTTCCGTCCT　246

Qy　5293　GGTCAGGATCTCTAGATTGGTCTCCGAGCTCCTCCTCCGTACTCCTCGTGTGCCTGTTCCTCT　5352
　　　　　|||||||||||| ||| |||||| ||||| ||||||||||||||| |||| ||| |||
Db　247　AGTCAGGATCTCT-GGTCCGTCAGCTTCTGCTACTCAGCCTGCTACTCCTCTCGTCCTCTCT　305

Qy　5353　CCTGTCCAGCTGCGCCACTGTGTGCTTCGTTCGCTTCCAGCTGTGGTCCACATTCTTCAGGAT　5412
　　　　　|||||||||||| |||||||||||||| ||||||| |||||| ||||||||| ||||||
Db　306　CTCCGTCCAGCTGCACCCTGTGTGGTGCCTCATTCCAGCCGTGGTCCGTGGTCCAAATTCTCTGTGAA　365

Qy　5413　TCTCTGAAAGTTAACCAGGTGAGAAGTGTTCCCGTGAGCAGCAGATCAGATTCT-C　5471
　　　　　||||| |||||| ||||| ||| |||||||| ||||||| ||| ||||||||| |
Db　366　AAGATTAACCGGGTGAGAATGCCCCCAGTTCCCCTGTCGCCCAGCTGCCCAGCTGCCCAGTCATGATTTTCC　425

Qy　5472　CCGGAAGTCAGGCTTCAGCCTTCCTCTTTCTCTGCCC　5531
　　　　　||-||| |||||| ||||| ||-||||||| ||||||
Db　426　CCAGAAGCCAGACTTCCAGCACCCGC--CCTCTGCCCAGCAACCTGGCACTCTTAGCAAA　483

Qy　5532　CCTCAGGCACCCTTACCCCACATAGACC　5559
　　　　　||-||| |||| |||||||||||||||
Db　484　CTTCAGCCCACCCTTCCCCCACATAGACC　511

RESULT 9
AZ407663
LOCUS　　　AZ407663　　　370 bp　　DNA　　linear　　GSS 03-OCT-2000
DEFINITION　1M0178C01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
　　　　　clone UUGC1M0178C01 R, genomic survey sequence.
ACCESSION　AZ407663
VERSION　　AZ407663.1　GI:10531676
KEYWORDS　GSS.
SOURCE　　Mus musculus (house mouse)
ORGANISM　Mus musculus

Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAAGGATCCAAGACTCGAGTTTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAAGGATCCAAGACTCAATTAATTAATTAACCCCCCCCCC 3'].
cDNA was cleaved with BamHI and XhoI. Vector:
pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match　　　4.1%;　　Score 236;　DB 5;　Length 447;
Best Local Similarity　93.2%;　Pred. No. 9.9e-51;
Matches　261;　Conservative　0;　Mismatches　17;　Indels　2;　Gaps　2;

Qy　1362　GAAAGGCCAGTCAATCAGATGCAGTCAGAAGAGAGGAGGAAGCCAACACACAGCTTGAGCA　1421
　　　　　|||| ||| ||||||||||| |||||||||||||||||||||| |||||||||||||||
Db　447　GANNGGCCANTCAATCAGATGCAGTCAGAGAGAGGATGGAAA-CCAACACACAGNTNGAGCA　389

Qy　1422　GAGGAAACAGAAAAAGGAGGAGATTCTGGGAGTAGAAGGAGGCCACAGAAGAAGAGCCCAGG　1481
　　　　　||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db　388　GAGGAAACAGAAAAAGGAGGAGAGNTTNTGGGAGTAGAAGGAGGCCAAAGAAGAAGAGCCCAGG　329

Qy　1482　CCCCCAAGTCTCCTTTATACCCTACATCCCGTCCCCAATTAAGCCACTCTTCTTCC　1541
　　　　　||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db　328　TCCCCCAAGTNTCNTNTTTATACCCTCATCCCGTCCTTCCAATTAAGCCANTTTTCTTCN　269

Qy　1542　TAGATCAGACCTGAGCTGCAGGGAAGAGACCCGTAGGAGGATCACACTGAGCACTGAAGGAG　1601
　　　　　||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db　268　TAGATCAGACCTGAGCTGCAGCGAAGAGACCCGTAGGAGGATCACACTGGGATGAAGGAG　209

Qy　1602　ATGTGTGAGAGAGTCCAGGGCAACCTAAGAGCCAGAGCCT　1641
　　　　　|||||||||||||||||||||| ||||||||||||||||
Db　208　ATGTGTGAGAGAGTCCAGGG-AACCTAAGAGCCAGAGCCT　170

RESULT 8
BF525096
LOCUS　　　BF525096　　　616 bp　　mRNA　　linear　　EST 11-DEC-2000
DEFINITION　UI-R-AC0-yh-h-03-0-UI.r1 UI-R-AC0 Rattus norvegicus cDNA clone
　　　　　UI-R-AC0-yh-h-03-0-UI 5', mRNA sequence.
ACCESSION　BF525096
VERSION　　BF525096.1　GI:11633063
KEYWORDS　EST.
SOURCE　　Rattus norvegicus (Norway rat)
ORGANISM　Rattus norvegicus
　　　　　Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
　　　　　Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
　　　　　Rattus.
REFERENCE　1 (bases 1 to 616)
AUTHORS　Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE　　Normalization and subtraction: two approaches to facilitate gene
　　　　　discovery
JOURNAL　Genome Res. 6 (9), 791-806 (1996)
MEDLINE　97044477
PUBMED　8889548
COMMENT　Contact: Soares, MB
　　　　　Coordinated Laboratory for Computational Genomics
　　　　　University of Iowa
　　　　　375 Newton Road , 4156　MEBRF, Iowa City, IA 52242, USA
　　　　　Tel: 319 335 8250
　　　　　Fax: 319 335 9565
　　　　　Email: bento-soares@uiowa.edu
　　　　　cDNA Library Preparation: M.B. Soares Lab Clone distribution:
　　　　　clones will be available through Research Genetics (www.resgen.com)
　　　　　This clone is also available through the I.M.A.G.E. Consortium at
　　　　　LLNL (info@image.llnl.gov). IMAGE ID= 1795343 The following
　　　　　repetitive elements were found in this cDNA sequence: 91-195,
　　　　　>BC1_MM#scRNA 124-183, >ID5#SINE/ID

```
LOCUS       BY000731                  390 bp    mRNA    linear   EST 06-DEC-2002
DEFINITION  BY000731 RIKEN full-length enriched, adult male heart Mus musculus
            cDNA clone 1020013I03 5', mRNA sequence.
ACCESSION   BY000731
VERSION     BY000731.1  GI:26060980
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 390)
  AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
  MEDLINE   22354683
   PUBMED   12466851
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,N., Sasaki,D.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences Mamm. Genome 12, 673-677 (2001)
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            RIKEN integrated sequence analysis (RISA) system-384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
```

```
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error:  0.00
            Plate: 0178  row: C  column: 01
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 370.
FEATURES             Location/Qualifiers
     source          1..370
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0178C01"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /note="Vector: PWD42rnv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1|, a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
```

```
ORIGIN

Query Match        2.5%;    Score 145.4;  DB 8;  Length 370;
Best Local Similarity  96.1%;  Pred. No. 1.1e-26;
Matches 149;  Conservative  0;  Mismatches  6;  Indels  0;  Gaps  0;

Qy     4062 GGCCTGGGTAGGGGAGGTGGTGTGAGACGCTCCTGCTCCTCCCGTCATGCCCTGAGCGCC 4121
            ||||||||||||||||||||||||||||||||||||||||||| |||| |||||||||||
Db        1 GGCCTGGGTAGGGGAGGTGGTGTGAGACGCTCCTGCTCCTCCCTATCTGCCCATCCGGCC   60

Qy     4122 CTTTGGGAGGAGGAAATGCGCCAAGGACTAAAAAAAGGCCATGGAGCCAGAGGGGCGAG 4181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 CTTTGGGAGGAGGAAATGCGCCAAGGACTAAAAAAAGGCCATGGAGCCAGAGGGGCCAG 120

Qy     4182 GGCAACAGACCTTTCATGGGCAAACCTTGGGCGCC 4216
            ||||||||||||||||||||||||||||||||||||
Db      121 GGCAACAGACCTTTCATGGGCAAACCTTGGGCGCC 155

RESULT 10
BY000731
```

Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.

FEATURES   Location/Qualifiers
     source   1..390
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="1020011103"
                     /sex="male"
                     /tissue_type="heart"
                     /dev_stage="adult"
                     /lab_host="BNN12"
                     /clone_lib="RIKEN full-length enriched, adult male heart"
                     /note="Site 1: SalI; Site 2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'.
                     GAGAGAGAGAGCCGGCCGCAACTCGAGTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. Second strand cDNA was prepared with the
                     primer adapter of sequence [5'
                     GAGAGAGAGAGGATCCAAGAGCTCAATTAATTAATTAAACCCCCCCCCCC 3'].
                     cDNA was cleaved with BamHI and XhoI. Vector: a modified
                     pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match   2.3%;   Score 133.4;   DB 5;   Length 390;
Best Local Similarity 95.8%;   Pred. No. 1.7e-23;
Matches 137;   Conservative   0;   Mismatches   6;   Indels   0;   Gaps   0;

Qy   5289 TCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCCTACTCCTCCTTCCTGCCTCGTTC 5348
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   10 TCCAACCCAGGATCTCTGATTCGTCTCCCAGCCTCCTACTCCTCTTCCTCCTGCCTGTTC 69

Qy   5349 CTCTCTCCGTCCAGCTGCCCACTGTGGTGCCTCGTTCCAGCTGGTCCACATTCTTCA 5408
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   70 CTCTCTCCGTCCAGCTGCCCACTGTGGTGCCTCGTTCCAGCTGGTCCACATTCTTCA 129

Qy   5409 GGATTCTCTGAAAAGTTAACCAG 5431
                  |||||||||||||||||||||||
Db   130 GGATTCTCTGAAAAGTTAACCAG 152

RESULT 11
BY061731
LOCUS   BY061731   409 bp   mRNA   linear   EST 06-DEC-2002
DEFINITION   BY061731 RIKEN full-length enriched, pooled tissues, 16 days
             embryo, etc. Mus musculus cDNA clone I920018N06 5', mRNA sequence.
ACCESSION   BY061731
VERSION   BY061731.1   GI:26181469
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 409)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
             Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
             Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
             Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
             Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
             Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
             Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
             Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
             Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
             Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
             Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
             Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
             Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,

Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
             of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
             Laboratory for Genome Exploration Research Group, RIKEN Genomic
             Sciences Center(GSC), Yokohama Institute
             The Institute of Physical and Chemical Research (RIKEN)
             1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
             Tel: 81-45-503-9222
             Fax: 81-45-503-9216
             Email: genome-res@gsc.riken.jp. URL:http://genome.gsc.riken.jp/
             Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
             Hirozane,T., Imotani,K., Ishii,Y., Itoh.M., Kawai,J., Konno,H.,
             Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
             Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
             Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
             Hayashizaki,Y. Direct Submission
             Computational Analysis of Full-Length Mouse cDNAs Compared with
             Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
             Normalization and subtraction of cap-trapper-selected cDNAs to
             prepare full-length cDNA libraries for rapid discovery of new
             genes. Genome Res. 10 (10), 1617-1630 (2000)
             RIKEN integrated sequence analysis (RISA) system--384-format
             sequencing pipeline with 384 multicapillary sequencer. Genome Res.
             10 (11), 1757-1771 (2000)
             Computer-based methods for the mouse full-length cDNA
             encyclopedia: real-time sequence clustering for construction of a
             nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
             cDNA library was prepared and sequenced in Mouse Genome
             Encyclopedia Project of Genome Exploration Research Group in Riken
             Genomic Sciences Center and Genome Science Laboratory in RIKEN.
             Division of Experimental Animal Research in Riken contributed to
             prepare mouse tissues.
             Please visit our web site (http://genome.gsc.riken.go.jp) for
             further details.
FEATURES   Location/Qualifiers
     source   1..409
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="I920018N06"
                     /clone_lib="RIKEN full-length enriched, pooled tissues, 16
                     days embryo, etc."
                     /note="pooled tissues: (dev_stage=16 days
                     embryo,tissue_type=heart,sex=mix),(dev_stage=16 days
                     embryo,tissue_type=kidney,sex=mix),(dev_stage=17 days
                     embryo,tissue_type=heart,sex=mix),(dev_stage=17 days
                     embryo,tissue_type=stomach,sex=mix),(dev_stage=17 days
                     embryo,tissue_type=kidney,sex=mix),(dev_stage=17 days
                     pregnant, adult,tissue_type=amnion,sex=female),
                     (dev_stage=13 days embryo,tissue_type=liver,sex=mix)"

ORIGIN

Query Match   2.3%;   Score 133.4;   DB 5;   Length 409;
Best Local Similarity 95.8%;   Pred. No. 1.7e-23;
Matches 137;   Conservative   0;   Mismatches   6;   Indels   0;   Gaps   0;

```
Qy    5289  TCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTTCCTCCTGCCTGTTC 5348
                 ||| ||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db      15  TCCAACCCAGGATCTCTGGATTGGTCTCCCAGCCTCTGCTACTCCTCCTTCCTGCCTGTTC 74

Qy    5349  CTCTCTCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGGTCCACATTCTTCA 5408
                 |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      75  CTCTCTCCCGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGGTCCACATTCTTCA 134

Qy    5409  GGATTCCTGAAAAGTTAACCAG 5431
            ||||||||||||||||||||||
Db     135  GGATTCCTGAAAAGTTAACCAG 157
```

RESULT 12

BY083328

LOCUS   BY083328   415 bp   mRNA   linear   EST 07-DEC-2002
DEFINITION BY083328 RIKEN full-length enriched, pooled tissues, adult spleen,
            etc. Mus musculus cDNA clone K630043E16 5', mRNA sequence.
ACCESSION   BY083328
VERSION   BY083328.1 GI:26193492
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 415)
AUTHORS   Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
            Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
            Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,
            Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
            Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
            Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,
            Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
            Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
            Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
            Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
            Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
            Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
            Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
            Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G.,
            Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
            Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
            Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,
            Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
            Verardo,R., Wagner,L., Wahlested,C., Wang,Y., Watanabe,Y.,
            Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
            Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
            Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
            Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
            Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
            Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
            Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
            Rogers,J., Birney,E. and Hayashizaki,Y.
TITLE   Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
JOURNAL   Nature 420, 563-573 (2002)
MEDLINE   22354683
PUBMED   12466851
COMMENT   Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
            Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S.,
            Hirozane,T., Imotani,K., Ishii,Y., Itoh,M., Kawai,J., Konno,H.,
            Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R.,
            Ohno,M., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K.,
            Shiraki,T., Tagami,M., Waki,K., Watahiki,A., Muramatsu,M. and
            Hayashizaki,Y. Direct Submission
            Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing Pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
```

FEATURES             Location/Qualifiers
     source          1..415
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="K630043E16"
                     /clone_lib="RIKEN full-length enriched, pooled tissues,
                     adult spleen, etc."
                     /note="(dev_stage=adult,tissue_type=spleen,sex=male),
                     (dev_stage=adult,tissue_type=kidney,sex=male),
                     (dev_stage=adult,tissue_type=testis,sex=male),
                     (dev_stage=adult,tissue_type=thymus,sex=male),
                     (dev_stage=adult,tissue_type=heart,sex=male),
                     (dev_stage=adult,tissue_type=colon,sex=male),
                     (dev_stage=adult,tissue_type=stomach,sex=male),
                     (dev_stage=adult,tissue_type=liver,sex=male),
                     (dev_stage=13 days embryo,tissue_type=whole
                     body,sex=mix),(dev_stage=14 days embryo,tissue_type=whole
                     body,sex=mix),(dev_stage=16 days embryo,tissue_type=whole
                     body,sex=mix),(dev_stage=17 days embryo,tissue_type=whole
                     body,sex=mix),(dev_stage=15 days pregnant,
                     adult,tissue_type=amnion,sex=female),(dev_stage=10 days
                     neonate,tissue_type=brain,sex=mix),(dev_stage=10 days
                     neonate,tissue_type=thymus,sex=mix),(dev_stage=10 days
                     neonate,tissue_type=heart,sex=mix)"

ORIGIN

Query Match         2.3%;  Score 133.4;  DB 5;  Length 415;
Best Local Similarity   95.8%;  Pred. No. 1.7e-23;
Matches   137;  Conservative   0;  Mismatches   6;  Indels   0;  Gaps   0;

```
Qy    5289  TCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTTCCTCCTGCCTGTTC 5348
                 |||||||||||| |||||||||||||||||||||||||||||||||| ||||||||
Db      15  TCCAACCCAGGATCTCTGGATTGGTCTCCCAGCCTCTGCTACTCCTCCTTCCTGCCTGTTC 74

Qy    5349  CTCTCTCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGGTCCACATTCTTCA 5408
                 |||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db      75  CTCTCTCCCGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGGTCCACATTCTTCA 134

Qy    5409  GGATTCCTGAAAAGTTAACCAG 5431
            ||||||||||||||||||||||
Db     135  GGATTCCTGAAAAGTTAACCAG 157
```

RESULT 13

CB596657

LOCUS   CB596657   739 bp   mRNA   linear   EST 03-APR-2003
DEFINITION AGENCOURT 12990591 NIH_MGC_178 Mus musculus cDNA clone
            IMAGE:30299997 5', mRNA sequence.
ACCESSION   CB596657
VERSION   CB596657.1 GI:29514513
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE   1 (bases 1 to 739)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: NDCM64 row: 1 column: 18
             High quality sequence stop: 590.
FEATURES   Location/Qualifiers
     source   1..739
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30299897"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_178"
                     /note="Organ: lung and heart; Vector: pDNR-LIB; Site_1:
                     SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); cDNA
                     made by oligo-dT priming and directionally cloned. 5' and
                     3' adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.5 kb
                     size fraction. Library created in the laboratory of M.
                     Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN

Query Match   2.3%; Score 133.4; DB 6; Length 739;
Best Local Similarity 95.8%; Pred. No. 2.1e-23;
Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy   5289 TCCTGGTCAGGATCTCTAGATTGGTCTCCAGCCTCTGCTACTCCTCTTCCTGCCTGTTC 5348
               || ||||||||||||||||| ||||||||||||||||||||||| |||||||||||||||
Db   17 TCCAACCCAGGATCTCTGGATTGGTCTCCAGCCTCTGCTACTCCTCTTCCTGCCTGTTC 76

Qy   5349 CTCTCTCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGTGGTCCACATTCTTCA 5408
          |||| | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CTCTCTCCGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGTGGTCCACATTCTTCA 136

Qy   5409 GGATTCTCTGAAAAGTTAACCAG 5431
          |||||||||||||||||||||||
Db   137 GGATTCTCTGAAAAGTTAACCAG 159

RESULT 14
CB599099
LOCUS   CB599099   741 bp   mRNA   linear   EST 03-APR-2003
DEFINITION AGENCOURT 13002811 NIH_MGC_178 Mus musculus cDNA clone
            IMAGE:30298800 5', mRNA sequence.
ACCESSION   CB599099
VERSION   CB599099.1 GI:29516955
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 741)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
            cDNA Library Preparation: Michael Brownstein Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM61 row: o column: 01
High quality sequence stop: 502.
FEATURES   Location/Qualifiers
     source   1..741
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30298800"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_178"
                     /note="Organ: lung and heart; Vector: pDNR-LIB; Site_1:
                     SfiI (ggccattatggcc); Site_2: SfiI (ggccgcctcggcc); cDNA
                     made by oligo-dT priming and directionally cloned. 5' and
                     3' adaptors were used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'. Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 0.5 kb
                     size fraction. Library created in the laboratory of M.
                     Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."
ORIGIN

Query Match   2.3%; Score 133.4; DB 6; Length 741;
Best Local Similarity 95.8%; Pred. No. 2.1e-23;
Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy   5289 TCCTGGTCAGGATCTCTAGATTGGTCTCCAGCCTCTGCTACTCCTCTTCCTGCCTGTTC 5348
               || ||||||||||||||||| ||||||||||||||||||||||| |||||||||||||
Db   17 TCCAACCCAGGATCTCTGGATTGGTCTCCAGCCTCTGCTACTCCTCTTCCTGCCTGTTC 76

Qy   5349 CTCTCTCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGTGGTCCACATTCTTCA 5408
          |||| | | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   77 CTCTCTCCGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGTGGTCCACATTCTTCA 136

Qy   5409 GGATTCTCTGAAAAGTTAACCAG 5431
          |||||||||||||||||||||||
Db   137 GGATTCTCTGAAAAGTTAACCAG 159

RESULT 15
CA490238
LOCUS   CA490238   783 bp   mRNA   linear   EST 09-MAR-2004
DEFINITION AGENCOURT_10825991 NIH_MGC_156 Mus musculus cDNA clone
            IMAGE:6755570 5', mRNA sequence.
ACCESSION   CA490238
VERSION   CA490238.1 GI:24953042
KEYWORDS   EST.
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 783)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Bradfield Laboratory
            cDNA Library Preparation: Mark Bittinger
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLKM00109 row: 1 column: 01
            High quality sequence stop: 543.
FEATURES   Location/Qualifiers
     source   1..783
                     /organism="Mus musculus"

```
          /mol_type="mRNA"
          /db_xref="taxon:10090"
          /clone="IMAGE:6755570"
          /lab_host="DH10B (phage-resistant)"
          /clone lib="NIH_MGC_156"
          /note="Organ: Heart: TCCD (DMSO vehicle) treated 48 hours
          IP injection; Vector: pDONR201; Site_1: attP2; Site_2:
          attP1; cDNA made by oligo-dT with attB2 site and
          directionally cloned. Priming sequence:
          5'-TTTCCTGCAGGCCGGCCACCACTTTGTACAAGAAAGCTGGGTTTTTTTTTTTTT
          TTTTT-3'. Full-length enriched library was constructed
          using the GeneRacer kit by Invitrogen, library
          amplification 16 cycles. Library constucted by Mark
          Bittinger in the Bradfield laboratory (McArdle Laboratory
          for Cancer Research, University of Wisconsin). Note: this
          is a NIH_MGC Library."

ORIGIN

Query Match      2.3%;   Score 133.4;  DB 6;   Length 783;
Best Local Similarity   95.8%;   Pred. No. 2.2e-23;
Matches  137;  Conservative    0;  Mismatches    6;   Indels    0;   Gaps    0;

Qy    5289  TCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCTTCCTGCCTGTTC  5348
            |||||| ||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db      51  TCCAACCCAGGATCTCTGGATTGGTCTCCCAGCCTCTGCTACTCCTCTTCCTGCCTGTTC   110

Qy    5349  CTCTCTCTGTCCAGCTGGCCACTGTGGTGCCTCGTTCCAGCTGTGGTCCACATTCTTCA  5408
            |||||| ||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     111  CTCTCTCCGTCCAGCTGCGCCACTGTGGTGCCTCGTTCCAGCTGTGGTCCACATTCTTCA   170

Qy    5409  GGATTCTCTGAAAAGTTAACCAG  5431
            |||||||||||||||||||||||
Db     171  GGATTCTCTGAAAAGTTAACCAG   193


Search completed: April 25, 2005, 14:31:21
Job time : 12067.2 secs
```

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

```
        GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.
```

Run on: April 25, 2005, 02:43:47 ; Search time 636.574 Seconds
                                           (without alignments)
                                           14741.485 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatcctgcaaggtcacaca........ccataaggttgagtcgac 5735

Scoring table: IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents NA:*
              1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
              2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
              3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
              4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
              5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
              6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                          SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1279.8	22.3	1679	2	US-08-365-486A-11	Sequence 11, Appl
	2	1279.8	22.3	1679	3	US-08-880-342-11	Sequence 11, Appl
	3	312.6	5.5	10728	4	US-09-376-774-5	Sequence 5, Appli
	4	311.8	5.4	520	1	US-08-485-971-7	Sequence 7, Appli
	5	311.8	5.4	520	1	US-08-275-876-7	Sequence 7, Appli
	6	311.8	5.4	520	1	US-08-383-754-7	Sequence 7, Appli
	7	311.8	5.4	520	2	US-08-485-978-7	Sequence 7, Appli
	8	311.8	5.4	520	2	US-08-486-814-7	Sequence 7, Appli
	9	311.8	5.4	520	2	US-08-487-472-7	Sequence 7, Appli
	10	311.8	5.4	520	3	US-08-485-740-7	Sequence 7, Appli
	11	311.8	5.4	520	3	US-09-162-184-7	Sequence 7, Appli
	12	311.8	5.4	520	3	US-09-161-902-7	Sequence 7, Appli
	13	311.8	5.4	520	3	US-09-489-777A-7	Sequence 7, Appli
	14	311.8	5.4	520	5	PCT-US95-08179-7	Sequence 7, Appli
c	15	310.4	5.4	450	1	US-08-076-726-13	Sequence 13, Appl
c	16	310.4	5.4	450	1	US-08-485-971-9	Sequence 9, Appli
c	17	310.4	5.4	450	1	US-08-260-452-6	Sequence 6, Appli
c	18	310.4	5.4	450	1	US-08-275-876-9	Sequence 9, Appli
c	19	310.4	5.4	450	1	US-08-383-754-9	Sequence 9, Appli
c	20	310.4	5.4	450	1	US-08-485-978-9	Sequence 9, Appli
c	21	310.4	5.4	450	2	US-08-481-970-6	Sequence 6, Appli
c	22	310.4	5.4	450	2	US-08-486-814-9	Sequence 9, Appli
c	23	310.4	5.4	450	2	US-08-487-472-9	Sequence 9, Appli
c	24	310.4	5.4	450	2	US-08-897-719-6	Sequence 6, Appli
c	25	310.4	5.4	450	3	US-08-485-740-9	Sequence 9, Appli
c	26	310.4	5.4	450	3	US-09-162-184-9	Sequence 9, Appli
c	27	310.4	5.4	450	3	US-09-161-902-9	Sequence 9, Appli
c	28	310.4	5.4	450	3	US-09-163-269-6	Sequence 6, Appli
c	29	310.4	5.4	450	3	US-09-489-777A-9	Sequence 9, Appli
c	30	310.4	5.4	450	4	US-09-281-674-6	Sequence 6, Appli
c	31	310.4	5.4	450	5	PCT-US95-08179-9	Sequence 9, Appl
	32	309.4	5.4	450	1	US-08-076-726-12	Sequence 12, Appl
	33	309.4	5.4	450	1	US-08-485-971-8	Sequence 8, Appli
	34	309.4	5.4	450	1	US-08-260-452-5	Sequence 5, Appli
	35	309.4	5.4	450	1	US-08-275-876-8	Sequence 8, Appli
	36	309.4	5.4	450	1	US-08-383-754-8	Sequence 8, Appli
	37	309.4	5.4	450	1	US-08-485-978-8	Sequence 8, Appli
	38	309.4	5.4	450	2	US-08-481-970-5	Sequence 5, Appli
	39	309.4	5.4	450	2	US-08-486-814-8	Sequence 8, Appli
	40	309.4	5.4	450	2	US-08-487-472-8	Sequence 8, Appli
	41	309.4	5.4	450	2	US-08-897-719-5	Sequence 5, Appli
	42	309.4	5.4	450	3	US-08-485-740-8	Sequence 8, Appli
	43	309.4	5.4	450	3	US-09-162-184-8	Sequence 8, Appli
	44	309.4	5.4	450	3	US-09-161-902-8	Sequence 8, Appli
	45	309.4	5.4	450	3	US-09-163-269-5	Sequence 5, Appli

                    ALIGNMENTS

RESULT 1
US-08-365-486A-11
; Sequence 11, Application US/0835486A
; Patent No. 5834306
; GENERAL INFORMATION:
; APPLICANT: Webster, Keith A.
; APPLICANT: Bishopric, Nanette H.
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
; TITLE OF INVENTION: Therapeutic Constructs
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Dehlinger & Associates
;   STREET: 350 Cambridge Avenue, Suite 250
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94306
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/365,486A
;   FILING DATE: 23-DEC-1994
;   CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
;   NAME: Sholtz, Charles K.
;   REGISTRATION NUMBER: 38,615
;   REFERENCE/DOCKET NUMBER: 8255-0018
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (415) 324-0880
;   TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1679 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;   INDIVIDUAL ISOLATE: Mouse alpha MHC promoter fragment
US-08-365-486A-11

Query Match   22.3%; Score 1279.8; DB 2; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;
```

```
Qy  3767 GAATTCTCTTACTATCAAAGGAAACTGAGTCGTGCACCTGCGTCTCCCT 3826
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GAATTCTCTTACTATCAAAGGAAACTGAGTCGTGCACCTGCGTCTCCCT 60

Qy  3827 AGACATCATGACTTTGTCTCTGGGAGCCAGACACTGTGAACTTCAGGTCT 3886
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGACATCATGACTTTGTCTCTGGGAGCCAGACACTGTGAACTTCAGGTCT 120

Qy  3887 GGAGGCTCCCCAGCCTGAAGCTATGCAGATAGCAGGGGGAAGGGGAGGAG 3946
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGAGGCTCCCCAGCCTGAAGCTATGCAGATAGCAGGGGGAAGGGGAGGAG 180

Qy  3947 CCTGGGATGGGAGCTTGTGTGTTGGAGGCAGACAGATATTAAGCCTGGAGGCAGAGAAGG 4006
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGGGATGGGAGCTTGTGTGTTGGAGGCAGACAGATATTAAGCCTGGAGGCAGAGAAGG 240

Qy  4007 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCT 4066
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCT 300

Qy  4067 GGGTAGGGGAGGTGGTGTGAGACGGCTCCTGTCTCCTGCGTATGGAGCCCTGAGGCCCTTG 4126
         ||||||||||||||||||||||||||||| |||| ||||||||
Db   301 GGGTAGGGGAGGTGGTGTGAGACGGCTCCTGTCTCCTCCTATCTGCCCATCCGGCCCTTG 360

Qy  4127 GGGAGGAGAATGTGCCCAAGGACTAAAAAAGGCCATGGAGCCAGAGGGGCGAGGGCAA 4186
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGGAGGAGAATGTGCCCAAGGACTAAAAAAGGCCATGGAGCCAGAGGGGCGAGGGCAA 420

Qy  4187 CAGACCTTTCATGGGCAAACCTTGGGGCCGTAGTGATCGATTGACAGAACTCGCCAAT 4246
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   421 CAGACCTTTCATGGGCAAACCTTGGGGCCCTGCTGTCCTGTCACCTCCAGACCCAG 480

Qy  4247 CGATACCCCTTCTTCTTCTCAACGACAGGAGGAACTCGAGTTTACCACTCCCTATCAGTG 4306
         |||||||||||
Db   481 GGATCAAAGGAGGAGGAGCCAGGACAGGAGGAAGTGGAGGAG--------------- 525

Qy  4307 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4366
         ||||||||||||
Db   526 -------------------------------------------------- 525

Qy  4367 CGAGTTTACCACTCCCTATCGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTA 4426
         ||||||||||||
Db   526 -------------------------------------------------- 525

Qy  4427 TCAGTGATGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGT 4486
         ||||||||||||
Db   526 -------------------------------------------------- 525

Qy  4487 GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4546
         ||||||||||||
Db   526 -------------------------------------------------- 525

Qy  4547 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGCTCGGTACCAGCAGAGGACTCCAAAT 4606
         ||||||||||||
Db   526 ---------------------GGTCCCAGCAGAGGACTCCAAAT 548

Qy  4607 TTAGGCAGCAGGTAGAGGTAAGAGGGCTGGAGGCTGGAGGCTGGAGAGCTGTCAGAG 4666
         |||||||||||||||||||||||||||||||
Db   549 TTAGGCAGCAGGTAGAGGTAAGAGGGCTGGAGGCTGGAGAGCTGTCAGAG 608

Qy  4667 ATTTCTCCAACCCAGGTAGAGGAGTTTCGGGTGGGGGGCTCTTCACCCACCACCAGACCT 4726
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   609 ATTTCTCCAACCCAGGTAGAGGAGTTTCGGGTGGGGGGCTCTTCACCCACCACCAGACCT 668

Qy  4727 CTCCCCACTCAGAAGGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCGGTTCAGAGATCT 4786
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   669 CTCCCCACTCAGAAGGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCGGTTCAGAGATCT 728

Qy  4787 GACAGGGTGGCCTTCCACCAGCCTCCAGGAAGTTCTCAGTGGCCAGGAGTTTCCACAAGAAA 4846
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   729 GACAGGGTGGCCTTCCACCAGCCTCCAGGAAGTTCTCAGTGGCCAGGAGTTTCCACAAGAAA 788

Qy  4847 CACTGGGATGCCCCTTCCCTTACGCTGTCTTCTCCCATCGTCTTCCTCTGGGGATGCTCCTCCC 4906
```

```
Db   789 CACTGGGATGCCCCTTCCCTTACGCTGTCTTCTCCCATCTTCCTCTGGGGATGCTCCTCCC 848
         |||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4907 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCTTCGTCTTGCCCTGTCGTCCCACTCCA 4966
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   849 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCTTCGTCTTGCCCTGTCGTCCCACTCCA 908

Qy  4967 TCTTTCTCTACTGTCTCCGTGCCTTCTTGCGTGCCTTCCTTCCACCCA 5026
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   909 TCTTTCTCTACTGTCTCCGTGCCTTCTTGCGTGCCTTCCTTCCACCCA 968

Qy  5027 TTTCTCACTTCACCTTTCTCCCCTTCTCATTTGTATTCATCCTTCCTTCCTTCCTTCCT 5086
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   969 TTTCTCACTTCACCTTTCTCCCCTTCTCATTTGTATTCATCCTTCCTTCCTTCCTTCCT 1028

Qy  5087 TCCTTCCTTCCTTCCTTCCTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 5146
         ||||||||||||||||||||||||||||||||||||||||||
Db  1029 TCCTTCCTTCCTTCCTTCCTCCTTCCTTCCTTCCTTCCTTCCCACCCTTATG 1088

Qy  5147 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGTTCCCACCCTTATG 5206
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1089 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGTTCCCACCCTTATG 1148

Qy  5207 TAAACAATCTTCCAGTGAGCCACCAGCTTCAGTGCTCTGGTCTGGGTGCTCCCAGCCTCTCA 5266
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1149 TAAACAATCTTCCAGTGAGCCACCAGCTTCAGTGCTGTGGTCTGGGTGCTCCCAGCCTCTCA 1208

Qy  5267 CCCCTGGCTGTCTGCTGTTCCATCCTGGTTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 5326
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1209 CCCCTGGCTGTCTGCTGTTCCATCCTGGTTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 1268

Qy  5327 CTACTCCCTCCTTCCTGCCTGGTCCTGTCCTCCTCTCTCCTCGTCTGTCCAGCTGGCCGTGTGCTCCGTTC 5386
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1269 CTACTCCCTCCTTCCTGCCTGGTCCTGTCCTCCTCTCTCCTCGTCTGTCCAGCTGGCCGTGTGCTCCGTTC 1328

Qy  5387 CAGCTGTGGTCCACATTCTTCAGGATTCCTGAAAAGTTAACCAGGTGAGAATGTTTCCC 5446
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1329 CAGCTGTGGTCCACATTCTTCAGGATTCCTGAAAAGTTAACCAGGTGAGAATGTTTCCC 1388

Qy  5447 CTGTAGACAGCAGATCACGATTCCCGGAAGTCAGGCTTCCAGCCTCTTCTTTCTCTGC 5506
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1389 CTGTAGACAGCAGATCACGATTCCCGGAAGTCAGGCTTCCAGCCCTCTTCTTTCTCTGC 1448

Qy  5507 CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGACACCCTTACCCCACCATAGACCTCTGACA 5566
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1449 CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACCATAGACCTCTGACA 1508

Qy  5567 GAGAAGCAGCAGGCACTTACATGAGTCCTGGTGGAGGAGCCATAGGCTACGGTGTAAAAGA 5626
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1509 GAGAAGCAGCAGGCACTTACATGAGTCCTGGTGGAGGAGCCATAGGCTACGGTGTAAAAGA 1568

Qy  5627 GGCAGGGAAGTGGTGGTGTGAGGAGAAGTCAGGACTTCACATGAGAAGCCTAGCCACACCAG 5686
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1569 GGCAGGGAAGTGGTGGTGTGAGGAGAAGTCAGGACTTCACATGAGAAGCCTAGCCACACCAG 1628

Qy  5687 AAATGACAGACAGATCCCTCCTATCCCCCATAAGAGTTTGAGT 5731
         |||||||||||||||||||||||||||||||||||||||||||
Db  1629 AAATGACAGACAGATCCCTCCTATCCCCCATAAGAGTTTGAGT 1673
```

```
CORRESPONDENCE ADDRESS:
;   ADDRESSEE:  Dehlinger & Associates
;   STREET:  350 Cambridge Avenue, Suite 250
;   CITY:  Palo Alto
;   STATE:  CA
;   COUNTRY:  USA
;   ZIP:  94306
COMPUTER READABLE FORM:
;   MEDIUM TYPE:  Floppy disk
;   COMPUTER:  IBM PC compatible
;   OPERATING SYSTEM:  PC-DOS/MS-DOS
;   SOFTWARE:  PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
;   APPLICATION NUMBER:  US/08/880,342
;   FILING DATE:  23-JUN-1997
;   CLASSIFICATION:  514
PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  PCT/IB95/00996
;   FILING DATE:  13-NOV-1995
PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:  US 08/365,486
;   FILING DATE:  23-DEC-1994
ATTORNEY/AGENT INFORMATION:
;   NAME:  Sholtz, Charles K.
;   REGISTRATION NUMBER:  38,615
;   REFERENCE/DOCKET NUMBER:  8255-0018.30
TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (415) 324-0880
;   TELEFAX:  (415) 324-0960
INFORMATION FOR SEQ ID NO: 11:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  1679 base pairs
;   TYPE:  nucleic acid
;   STRANDEDNESS:  double
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA (genomic)
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   ORIGINAL SOURCE:
;   INDIVIDUAL ISOLATE:  Mouse alpha MHC promoter fragment
US-08-880-342-11
```

```
Query Match      22.3%;   Score 1279.8;   DB 3;   Length 1679;
Best Local Similarity   82.2%;   Pred. No. 0;
Matches 1616;   Conservative   0;   Mismatches   57;   Indels  292;   Gaps    1;
```

```
Qy   3767 GAATTCTCTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCCT 3826
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAATTCTCTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCCT 60

Qy   3827 AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA 3886
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA 120

Qy   3887 GGAGGCTCCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAGGGGAGAG 3946
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGAGGCTCCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAGGGGAGAG 180

Qy   3947 CCTGGGGATGGGAGCTTGTCGTTGGAGGGACAGAGATATTAAGCCTGGAAGAGAAGG 4006
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGGGGATGGGAGCTTGTCGTTGGAGGGACAGAGATATTAAGCCTGGAAGAGAAGG 240

Qy   4007 TGACCCTTACCCAGTTGTTCAACTCACCTTCAGATTAAAAATAACTGAGGTAAGGGCCT 4066
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGACCCTTACCCAGTTGTTCAACTCACCTTCAGATTAAAAATAACTGAGGTAAGGGCCT 300

Qy   4067 GGGTAGGGGAGGTGGTGTGAGACGCTCCTGTCTCCTCCTCATCGCCCATCGGCCCTTTG 4126
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGTAGGGGAGGTGGTGTGAGACGCTCCTGTCTCCTCCTCATCGCCCATCGGCCCTTTG 360

Qy   4127 GGGAGGAGGAATGCCCAAGGACTAAAAAAAGGCCATGAGCCAGAGGGGGCGAGGGCAA 4186
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGGAGGAGGAATGCCCAAGGACTAAAAAAAGGCCATGAGCCAGAGGGGGCGAGGGCAA 420
```

```
Qy   4187 CAGACCTTCATGGGCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAAT 4246
               ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGACCTTCATGGGCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAAG 480

Qy   4247 CGATACCCTTCTTCTTCTAACGGACGAGGAGGGAACTCGAGTTTACCACTCCTATCAGTG 4306
               |||   ||||||||||||    |    |
Db    481 GGATCAAAGGAGGAGGAGCCAGGAGGAGGAAGTGGGGAGGGAG-------- 525

Qy   4307 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGT 4366
Db    526 ---------- 525

Qy   4367 CGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCTA 4426
Db    526 ---------- 525

Qy   4427 TCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGT 4486
Db    526 ---------- 525

Qy   4487 GAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4546
Db    526 ---------- 525

Qy   4547 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTCCGGTACCAGCAGAGACTCCAAAT 4606
                                             |||||||||||||||||||
Db    526 -------GGTCCCAGCAGAGAGACTCCAAAT 548

Qy   4607 TTAGGCAGGCATATGGGATGGGATATAAAGGGGCTGAGCACTGAGAGCTGTCAGAG 4666
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    549 TTAGGCAGGCATATGGGATGGGATATAAAGGGGCTGAGCACTGAGAGCTGTCAGAG 608

Qy   4667 ATTTCTCCAACCCAGGTAAGAGGGAGTTCGGGTGGGGGCTCTTCACCCACCACCAGACCT 4726
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    609 ATTTCTCCAACCCAGGTAAGAGGGAGTTCGGGTGGGGGCTCTTCACCCACCACCAGACCT 668

Qy   4727 CTCCCACCTAGAAGGAAACTGCCTTCCTGGAAGTTCCTTCCTGGAAGTTCAGGCCGGT 4786
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    669 CTCCCACCTAGAAGGAAACTGCCTTCCTGGAAGTTCCTTCCTGGAAGTTCAGGCCGGT 728

Qy   4787 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCCAGTGCCAGGAGGTTTCCACAGAAA 4846
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    729 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCCAGTGCCAGGAGGTTTCCACAGAAA 788

Qy   4847 CACTGGATGCCCCTTCCCTTACGCTGTCTTCCCATCTTCCTCCTCGGGGATGCTCCTCCC 4906
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    789 CACTGGATGCCCCTTCCCTTACGCTGTCTTCCCATCTTCCTCCTCGGGGATGCTCCTCCC 848

Qy   4907 CGTCTTGGTTTATCTGGCCTTCAGCAAGATTTGCCCTGCGTCTTCCTTTCCACTCCA 4966
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    849 CGTCTTGGTTTATCTGGCCTTCAGCAAGATTTGCCCTGCGTCTTCCTTTCCACTCCA 908

Qy   4967 TCTTTCTCACTGTCTCCGTGCCTTCGCCTTGCCTTGCGTCTTCCTTTCCTTTCCACCCA 5026
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    909 TCTTTCTCACTGTCTCCGTGCCTTCGCCTTGCCTTGCGTCTTCCTTTCCTTTCCACCCA 968

Qy   5027 TTTCTCACTTCACCTTTTCTCCCCTTCTCCCCTTCTCATTGTATTCATCCTTCCTTCCT 5086
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    969 TTTCTCACTTCACCTTTTCTCCCCTTCTCCCCTTCTCATTGTATTCATCCTTCCTTCCT 1028

Qy   5087 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 5146
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1029 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 1088

Qy   5147 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCTATG 5206
               ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   1089 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTATG 1148

Qy   5207 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGCTCGTGGTGCTCTCTTACCTTCCTCA 5266
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1149 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGCTCGTGGTGCTCTCTTACCTTCCTCA 1208
```

Qy 5267 CCCCTGGCTTGTCCTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 5326
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1209 CCCCTGGCTTGTCCTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 1268

Qy 5327 CTATCTCCTCTGCCTGTTCTCTCTCGTGCCACTGTGGTGCCTCGTTC 5386
          |||||||||||||||||||||||||||||||||||||||||||||||
Db 1269 CTATCTCCTCTGCGCTGTTCTCTCTCGTGCCACTGTGGTGCCTCGTTC 1328

Qy 5387 CAGCTGTGGTCCACATTCTTCAGGATTCTCTGAAAAAGTTAACCAGGTGAGAATGTTTCCC 5446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1329 CAGCTGTGGTCCACATTCTTCAGGATTCTCTGAAAAAGTTAACCAGGTGAGAATGTTTCCC 1388

Qy 5447 CTGTAGACAGCAGATCACGATTCCCCGGAAGTCAGGCTTCCAGCCCTCTCTTTCTCTGC 5506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1389 CTGTAGACAGCAGATCACGATTCCCCGGAAGTCAGGCTTCCAGCCCTCTCTTTCTCTGC 1448

Qy 5507 CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCCACATAGACCTCTGACA 5566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1449 CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCCACATAGACCTCTGACA 1508

Qy 5567 GAGAAGCAGGCACTTACATGGAGTCCTGGTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 5626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1509 GAGAAGCAGGCACTTACATGGAGTCCTGGTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 1568

Qy 5627 GGCAGGGAAGTGGTGGTGGTAGGAAAGTCAGGACTTCACATAGAGCCTAGCCCACACCAG 5686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1569 GGCAGGGAAGTGGTGGTGGTAGGAAAGTCAGGACTTCACATAGAGCCTAGCCCACACCAG 1628

Qy 5687 AAATGACAGACAGATCCCTCCCTATCTCCCCATAAGAGTTTGAGT 5731
          |||||||||||||||||||||||||||||||||||||||||||||
Db 1629 AAATGACAGACAGATCCCTCCCTATCTCCCCATAAGAGTTTGAGT 1673

RESULT 3
US-09-376-774-5
; Sequence 5, Application US/09376774
; Patent No. 6759236
; GENERAL INFORMATION:
;   APPLICANT: Fung, Yuen Kai
;   APPLICANT: Gomer, Charles
;   APPLICANT: T'Ang, Anne
;   TITLE OF INVENTION: Methods To Enhance And Confine Expression
;   TITLE OF INVENTION: Of Genes
;   FILE REFERENCE: D6087
;   CURRENT APPLICATION NUMBER: US/09/376,774
;   CURRENT FILING DATE: 2003-03-21
;   PRIOR APPLICATION NUMBER: 60/096,947
;   PRIOR FILING DATE: 1998-08-18
;   NUMBER OF SEQ ID NOS: 5
; SEQ ID NO 5
;   LENGTH: 10728
;   TYPE: DNA
;   ORGANISM: Unknown
;   FEATURE:
;   NAME/KEY: misc feature
;   OTHER INFORMATION: recombinant vector pDATH-TNF?
US-09-376-774-5

  Query Match   5.5%; Score 312.6; DB 4; Length 10728;
  Best Local Similarity 97.2%; Pred. No. 7.7e-82;
  Matches   318; Conservative   0; Mismatches   9; Indels   0; Gaps   0;

Qy 4265 AACCGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGT 4324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 212   ATCCGACAAGAAGCTTGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGT 271

Qy 4325 CGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTA 4384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 272   CGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTA 331

Qy 4385 TCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGT 4444
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 332   TCAGTGATGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAGT 391

Qy 4445 GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAGTTTACCAC 4504
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 392   GAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTCGAGTTTACCAC 451

Qy 4505 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 452   TCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 511

Qy 4565 AAAAGTGAAAGTCGAGCTCGGTACCAG 4591
          |||||||||||||||||||||||||||
Db 512   AAAAGTGAAAGTCGAGCTCGGTACCCG 538

RESULT 4
US-08-485-971-7
; Sequence 7, Application US/08485971
; Patent No. 5589362
; GENERAL INFORMATION:
;   APPLICANT: Bujard, Hermann
;   APPLICANT: Gossen, Manfred
;   APPLICANT: Hillen, Wolfgang
;   APPLICANT: Helbl, Vera
;   APPLICANT: Schnappinger, Dirk
;   TITLE OF INVENTION: Tetracycline-Regulated Transcriptional
;   NUMBER OF SEQUENCES: 28
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: LAHIVE & COCKFIELD
;   STREET: 60 State Street, Suite 510
;   CITY: Boston
;   STATE: Massachusetts
;   COUNTRY: USA
;   ZIP: 02109-1875
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/485,971
;   FILING DATE:
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/383,754
;   FILING DATE: 03-FEB-1995
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/270,637
;   FILING DATE: 01-JULY-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/260,452
;   FILING DATE: 14-JUNE-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/076,327
;   FILING DATE: 14-JUNE-1993
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/076,726
;   FILING DATE: 14-JUNE-1993
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: DeConti, Giulio A. Jr.
;   REGISTRATION NUMBER: 31,503
;   REFERENCE/DOCKET NUMBER: BBI-009CP7
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (617)227-7400
;   TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 520 base pairs
;   TYPE: nucleic acid

```
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA
US-08-485-971-7

Query Match             5.4%;  Score 311.8; DB 1;  Length 520;
Best Local Similarity  99.4%;  Pred. No. 1.7e-82;
  Matches 313;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATAGAGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 TCCCTATCAGTGATAGAGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGTCGGTACCAG 4591
          |||||||||||||
Db    363 CGAGTCGGTACCCG 377
```

```
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA
US-08-275-876-7

Query Match             5.4%;  Score 311.8; DB 1;  Length 520;
Best Local Similarity  99.4%;  Pred. No. 1.7e-82;
  Matches 313;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
          ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 GGATCTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATAGAGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 TCCCTATCAGTGATAGAGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGTCGGTACCAG 4591
          |||||||||||||
Db    363 CGAGTCGGTACCCG 377
```

```
RESULT 5
US-08-275-876-7
; Sequence 7, Application US/08275876
; Patent No. 5654168
; GENERAL INFORMATION:
;     APPLICANT: Bujard, Hermann
;     APPLICANT: Gossen, Manfred
;     TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
;     NUMBER OF SEQUENCES: 15
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: LAHIVE & COCKFIELD
;        STREET: 60 State Street, suite 510
;        CITY: Boston
;        STATE: Massachusetts
;        COUNTRY: USA
;        ZIP: 02109-1875
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: ASCII Text
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/275,876
;        FILING DATE:
;        CLASSIFICATION: 436
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US
;        FILING DATE: 01-JULY-94
;        CLASSIFICATION: 436
;     ATTORNEY/AGENT INFORMATION:
;        NAME: DeConti, Giulio A. Jr.
;        REGISTRATION NUMBER: 31,503
;        REFERENCE/DOCKET NUMBER: BBI-009CP
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (617)227-7400
;        TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 520 base pairs
```

```
RESULT 6
US-08-383-754-7
; Sequence 7, Application US/08383754
; Patent No. 5789156
; GENERAL INFORMATION:
;     APPLICANT: Bujard, Hermann
;     APPLICANT: Gossen, Manfred
;     TITLE OF INVENTION: Tetracycline-Regulated Transcriptional
;     TITLE OF INVENTION: Inhibitors
;     NUMBER OF SEQUENCES: 28
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: LAHIVE & COCKFIELD
;        STREET: 60 State Street, suite 510
;        CITY: Boston
;        STATE: Massachusetts
;        COUNTRY: USA
;        ZIP: 02109-1875
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: ASCII Text
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/383,754
;        FILING DATE:
;        CLASSIFICATION: 436
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/275,876
;        FILING DATE: 15-JULY-94
;        CLASSIFICATION: 436
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/270,637
;        FILING DATE: 01-JULY-94
;        CLASSIFICATION: 436
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/260,452
;        FILING DATE: 14-JUNE-94
;        CLASSIFICATION: 436
```

```
;        SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/485,978
;        FILING DATE:
;        CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/383,754
;        FILING DATE: 03-FEB-1995
;        CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/275,876
;        FILING DATE: 15-JULY-1994
;        CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/270,637
;        FILING DATE: 01-JULY-1994
;        CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/260,452
;        FILING DATE: 14-JUNE-1994
;        CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/076,327
;        FILING DATE: 14-JUNE-1993
;        CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/076,726
;        FILING DATE: 14-JUNE-1993
;        CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;        NAME: DeConti, Giulio A. Jr.
;        REGISTRATION NUMBER: 31,503
;        REFERENCE/DOCKET NUMBER: BBI-009CP6
;   TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (617)227-7400
;        TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;        LENGTH: 520 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: double
;        TOPOLOGY: linear
;        MOLECULE TYPE: DNA
US-08-485-978-7

Query Match          5.4%;     Score 311.8;  DB 1;  Length 520;
Best Local Similarity   99.4%;     Pred. No. 1.7e-82;
Matches  313;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 GGATCTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG 4396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 TCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAGAAAAGTGAAAGT 4576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCAG 4591
          ||||||||||||||
Db    363 CGAGCTCGGTACCCG 377
```

```
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/076,327
;        FILING DATE: 14-JUNE-93
;        CLASSIFICATION: 436
;   PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/076,726
;        FILING DATE: 14-JUNE-93
;        CLASSIFICATION: 436
;   ATTORNEY/AGENT INFORMATION:
;        NAME: DeConti, Giulio A. Jr.
;        REGISTRATION NUMBER: 31,503
;        REFERENCE/DOCKET NUMBER: BBI-009CP2
;   TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (617)227-7400
;        TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;        LENGTH: 520 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: double
;        TOPOLOGY: linear
;        MOLECULE TYPE: DNA
US-08-383-754-7

Query Match          5.4%;     Score 311.8;  DB 1;  Length 520;
Best Local Similarity   99.4%;     Pred. No. 1.7e-82;
Matches  313;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
          ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 GGATCTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG 4396
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 TCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAGAAAAGTGAAAGT 4576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCAG 4591
          ||||||||||||||
Db    363 CGAGCTCGGTACCCG 377

RESULT 7
US-08-485-978-7
; Sequence 7, Application US/08485978
; Patent No. 5814618
;   GENERAL INFORMATION:
;        APPLICANT: Bujard, Hermann
;        APPLICANT: Gossen, Manfred
;        TITLE OF INVENTION: Methods for Regulating Gene Expression
;        NUMBER OF SEQUENCES: 28
;   CORRESPONDENCE ADDRESS:
;        ADDRESSEE: LAHIVE & COCKFIELD
;        STREET: 60 State Street, Suite 510
;        CITY: Boston
;        STATE: Massachusetts
;        COUNTRY: USA
;        ZIP: 02109-1875
;   COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
RESULT 8
US-08-486-814-7
; Sequence 7, Application US/08486814
; Patent No. 5866755
; GENERAL INFORMATION:
;   APPLICANT:  Bujard, Hermann
;   APPLICANT:  Gossen, Manfred
;   TITLE OF INVENTION: Animals Transgenic for a Tetracycline-  Controlled Transcr
;   NUMBER OF SEQUENCES:  28
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: LAHIVE & COCKFIELD
;     STREET:  60 State Street, Suite 510
;     CITY:  Boston
;     STATE:  Massachusetts
;     COUNTRY:  USA
;     ZIP:  02109-1875
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/486,814
;     FILING DATE:
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/383,754
;     FILING DATE:  03-FEB-1995
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/275,876
;     FILING DATE:  15-JULY-1994
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/270,637
;     FILING DATE:  01-JULY-1994
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/260,452
;     FILING DATE:  14-JUNE-1994
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/076,327
;     FILING DATE:  14-JUNE-1993
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/076,726
;     FILING DATE:  14-JUNE-1993
;     CLASSIFICATION:  800
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  DeConti, Giulio A. Jr.
;     REGISTRATION NUMBER:  31,503
;     REFERENCE/DOCKET NUMBER:  BBI-009CP4
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (617)227-7400
;     TELEFAX:  (617)227-5941
;   INFORMATION FOR SEQ ID NO: 7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  520 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA
US-08-486-814-7

Query Match             5.4%;  Score 311.8;  DB 2;  Length 520;
Best Local Similarity   99.4%;  Pred. No. 1.7e-82;
Matches   313;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAGTGAAAGTCGAGTTTACCAC 4336
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 GGATCTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAGAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCTCCCTATCAGTGATAGAG 4396
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 TCCCTATCAGTGATGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCTCCCTATCAGTGATAGAG 182

Qy   4397 AAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGT 4456
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAGTGAGTTTACCACTCCCTATCAGTG 4516
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 TACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGT 4576
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 ATAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCACG 4591
          |||||||||||||||
Db    363 CGAGCTCGGTACCCG 377

RESULT 9
US-08-487-472-7
; Sequence 7, Application US/08487472
; Patent No. 5912411
; GENERAL INFORMATION:
;   APPLICANT:  Bujard, Hermann
;   APPLICANT:  Gossen, Manfred
;   TITLE OF INVENTION:  Animal Transgenic for a Tetracycline- Inducible Transcription
;   NUMBER OF SEQUENCES:  28
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  LAHIVE & COCKFIELD
;     STREET:  60 State Street, Suite 510
;     CITY:  Boston
;     STATE:  Massachusetts
;     COUNTRY:  USA
;     ZIP:  02109-1875
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE:  Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  ASCII Text
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/487,472
;     FILING DATE:
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/383,754
;     FILING DATE:  03-FEB-1995
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/275,876
;     FILING DATE:  15-JULY-1994
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/270,637
;     FILING DATE:  01-JULY-94
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/260,452
;     FILING DATE:  14-JUNE-1994
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/076,327
;     FILING DATE:  14-JUNE-1993
;     CLASSIFICATION:  800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  US 08/076,726
;     FILING DATE:  14-JUNE-1993
;     CLASSIFICATION:  800
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  DeConti, Giulio A. Jr.
```

```
;   REGISTRATION NUMBER: 31,503
;   REFERENCE/DOCKET NUMBER: BBI-009CP3
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE:  (617)227-7400
;   TELEFAX:  (617)227-5941
;   INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH:  520 base pairs
;   TYPE:  nucleic acid
;   STRANDEDNESS:  double
;   TOPOLOGY:  linear
;   MOLECULE TYPE:  DNA
US-08-487-472-7

Query Match      5.4%;     Score 311.8;  DB 2;  Length 520;
Best Local Similarity    99.4%;   Pred. No. 1.7e-82;
Matches  313;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy  4277  GGAACTCGAGTTTACCACTCCCTATCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCAC  4336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63  GGATCTCGAGTTTACCACTCCCTATCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCAC  122

Qy  4337  TCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG  4396
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123  TCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG  182

Qy  4397  AAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTT  4456
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   183  AAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTT  242

Qy  4457  TACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  4516
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   243  TACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  302

Qy  4517  ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGT  4576
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   303  ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGT  362

Qy  4577  CGAGCTCGGTACCAG  4591
          |||||||||||||
Db   363  CGAGCTCGGTACCCG  377

RESULT 10
US-08-485-740-7
;  Sequence 7, Application US/08485740
;  Patent No. 6004941
;  GENERAL INFORMATION:
;     APPLICANT: Bujard, Hermann
;     APPLICANT: Gossen, Manfred
;     TITLE OF INVENTION: Methods for Regulating Gene Expression
;     NUMBER OF SEQUENCES: 28
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: LAHIVE & COCKFIELD
;        STREET: 60 State Street, Suite 510
;        CITY: Boston
;        STATE: Massachusetts
;        COUNTRY: USA
;        ZIP: 02109-1875
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: ASCII Text
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/485,740
;        FILING DATE:
;        CLASSIFICATION: 435
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/383,754
;        FILING DATE: 03-FEB-1995
;        CLASSIFICATION: 435
;     PRIOR APPLICATION DATA:
```

```
;        APPLICATION NUMBER: US 08/275,876
;        FILING DATE: 15-JULY-1994
;        CLASSIFICATION: 435
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/270,637
;        FILING DATE: 01-JULY-1994
;        CLASSIFICATION: 435
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/260,452
;        FILING DATE: 14-JUNE-1994
;        CLASSIFICATION: 435
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/076,327
;        FILING DATE: 14-JUNE-1993
;        CLASSIFICATION: 435
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US 08/076,726
;        FILING DATE: 14-JUNE-1993
;        CLASSIFICATION: 435
;     ATTORNEY/AGENT INFORMATION:
;        NAME: DeConti, Giulio A. Jr.
;        REGISTRATION NUMBER: 31,503
;        REFERENCE/DOCKET NUMBER: BBI-009CP5
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  (617)227-7400
;        TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO: 7:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  520 base pairs
;        TYPE:  nucleic acid
;        STRANDEDNESS:  double
;        TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-08-485-740-7

Query Match      5.4%;     Score 311.8;  DB 3;  Length 520;
Best Local Similarity    99.4%;   Pred. No. 1.7e-82;
Matches  313;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy  4277  GGAACTCGAGTTTACCACTCCCTATCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCAC  4336
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63  GGATCTCGAGTTTACCACTCCCTATCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCAC  122

Qy  4337  TCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG  4396
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123  TCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAG  182

Qy  4397  AAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTT  4456
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   183  AAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTT  242

Qy  4457  TACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  4516
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   243  TACCACTCCCTATCAGTGATGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  302

Qy  4517  ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGT  4576
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   303  ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGAAAGT  362

Qy  4577  CGAGCTCGGTACCAG  4591
          |||||||||||||
Db   363  CGAGCTCGGTACCCG  377

RESULT 11
US-09-162-184-7
;  Sequence 7, Application US/09162184A
;  Patent No. 6136954
;  GENERAL INFORMATION:
;     APPLICANT: Bujard, Hermann
;     APPLICANT: Gossen, Manfred
;     TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
                          Activator Fusion Proteins
```

```
                NUMBER OF SEQUENCES: 37
                CORRESPONDENCE ADDRESS:
                   ADDRESSEE: LAHIVE & COCKFIELD
                   STREET: 28 State Street
                   CITY: Boston
                   STATE: Massachusetts
                   COUNTRY: USA
                   ZIP: 02109-1875
                COMPUTER READABLE FORM:
                   MEDIUM TYPE: Floppy disk
                   COMPUTER: IBM PC compatible
                   OPERATING SYSTEM: PC-DOS/MS-DOS
                   SOFTWARE: ASCII Text
                CURRENT APPLICATION DATA:
                   APPLICATION NUMBER: US/09/162,184A
                   FILING DATE: 28-Sep-1998
                   CLASSIFICATION: <Unknown>
                PRIOR APPLICATION DATA:
                   APPLICATION NUMBER: US 08/383,754
                   FILING DATE: 03-FEB-1995
                   APPLICATION NUMBER: US 08/275,876
                   FILING DATE: 15-JULY-1994
                   APPLICATION NUMBER: US 08/270,637
                   FILING DATE: 01-JULY-1994
                   APPLICATION NUMBER: US 08/260,452
                   FILING DATE: 14-JUNE-1994
                   APPLICATION NUMBER: US 08/076,327
                   FILING DATE: 14-JUNE-1993
                   APPLICATION NUMBER: US 08/076,726
                   FILING DATE: 14-JUNE-1993
                ATTORNEY/AGENT INFORMATION:
                   NAME: DeConti, Giulio A. Jr.
                   REGISTRATION NUMBER: 31,503
                   REFERENCE/DOCKET NUMBER: BBI-009C6CN
                TELECOMMUNICATION INFORMATION:
                   TELEPHONE: (617)227-7400
                   TELEFAX: (617)742-4214
                INFORMATION FOR SEQ ID NO: 7:
                   SEQUENCE CHARACTERISTICS:
                      LENGTH: 520 base pairs
                      TYPE: nucleic acid
                      STRANDEDNESS: double
                      TOPOLOGY: linear
                      MOLECULE TYPE: DNA
                   SEQUENCE DESCRIPTION: SEQ ID NO: 7:
               US-09-162-184-7
```

```
      Query Match        5.4%;  Score 311.8;  DB 3;  Length 520;
      Best Local Similarity  99.4%;  Pred. No. 1.7e-82;
      Matches 313;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy    4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy    4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy    4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy    4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 362

Qy    4577 CGAGCTCGGTACCAG 4591
                |||||||||||
```

```
Db     363 CGAGCTCGGTACCCG 377

      RESULT 12
      US-09-161-902-7
      ; Sequence 7, Application US/09161902
      ; Patent No. 6242667
      ; GENERAL INFORMATION:
           APPLICANT: Bujard, Hermann
           APPLICANT: Gossen, Manfred
           TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible Transcription
           NUMBER OF SEQUENCES: 28
           CORRESPONDENCE ADDRESS:
              ADDRESSEE: LAHIVE & COCKFIELD
              STREET: 60 State Street, Suite 510
              CITY: Boston
              STATE: Massachusetts
              COUNTRY: USA
              ZIP: 02109-1875
           COMPUTER READABLE FORM:
              MEDIUM TYPE: Floppy disk
              COMPUTER: IBM PC compatible
              OPERATING SYSTEM: PC-DOS/MS-DOS
              SOFTWARE: ASCII Text
           CURRENT APPLICATION DATA:
              APPLICATION NUMBER: US/09/161,902
              FILING DATE:
              CLASSIFICATION:
           PRIOR APPLICATION DATA:
              APPLICATION NUMBER: 08/487,472
              FILING DATE:
              CLASSIFICATION:
           PRIOR APPLICATION DATA:
              APPLICATION NUMBER: US 08/275,876
              FILING DATE: 15-JULY-1994
              CLASSIFICATION:
           PRIOR APPLICATION DATA:
              APPLICATION NUMBER: US 08/270,637
              FILING DATE: 01-JULY-94
              CLASSIFICATION:
           PRIOR APPLICATION DATA:
              APPLICATION NUMBER: US 08/260,452
              FILING DATE: 14-JUNE-1994
              CLASSIFICATION:
           PRIOR APPLICATION DATA:
              APPLICATION NUMBER: US 08/076,327
              FILING DATE: 14-JUNE-1993
              CLASSIFICATION:
           PRIOR APPLICATION DATA:
              APPLICATION NUMBER: US 08/076,726
              FILING DATE: 14-JUNE-1993
              CLASSIFICATION:
           ATTORNEY/AGENT INFORMATION:
              NAME: DeConti, Giulio A. Jr.
              REGISTRATION NUMBER: 31,503
              REFERENCE/DOCKET NUMBER: BBI-009CP3
           TELECOMMUNICATION INFORMATION:
              TELEPHONE: (617)227-7400
              TELEFAX: (617)227-5941
           INFORMATION FOR SEQ ID NO: 7:
              SEQUENCE CHARACTERISTICS:
                 LENGTH: 520 base pairs
                 TYPE: nucleic acid
                 STRANDEDNESS: double
                 TOPOLOGY: linear
                 MOLECULE TYPE: DNA
           ; US-09-161-902-7
```

```
      Query Match        5.4%;  Score 311.8;  DB 3;  Length 520;
      Best Local Similarity  99.4%;  Pred. No. 1.7e-82;
      Matches 313;  Conservative  0;  Mismatches  2;  Indels  2;  Gaps  0;

Qy    4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
```

```
Db    63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC  122
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy  4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG  4396
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG   182

Qy  4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT  4456
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT   242

Qy  4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  4516
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG   302

Qy  4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT  4576
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT   362

Qy  4577 CGAGCTCGGTACCAG  4591
         |||||||||||
Db   363 CGAGCTCGGTACCCG   377


RESULT 13
US-09-489-777A-7
; Sequence 7, Application US/09489777A
; Patent No. 6271348
; GENERAL INFORMATION:
   APPLICANT: Bujard, Hermann
            Gossen, Manfred
   TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
                      Inhibitor Fusion Proteins
   NUMBER OF SEQUENCES: 37
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: LAHIVE & COCKFIELD
      STREET: 28 State Street
      CITY: Boston
      STATE: Massachusetts
      COUNTRY: USA
      ZIP: 02109-1875
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: ASCII Text
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/489,777A
      FILING DATE: 24-Jan-2000
      CLASSIFICATION: <Unknown>
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 09/162,184
      FILING DATE: 28-SEP-1998
      APPLICATION NUMBER: US 08/485,978
      FILING DATE: 07-JUN-1995
      APPLICATION NUMBER: US 08/383,754
      FILING DATE: 03-FEB-1995
      APPLICATION NUMBER: US 08/275,876
      FILING DATE: 15-JULY-1994
      APPLICATION NUMBER: US 08/270,637
      FILING DATE: 01-JULY-1994
      APPLICATION NUMBER: US 08/260,452
      FILING DATE: 14-JUNE-1994
      APPLICATION NUMBER: US 08/076,327
      FILING DATE: 14-JUNE-1993
      APPLICATION NUMBER: US 08/076,726
      FILING DATE: 14-JUNE-1993
   ATTORNEY/AGENT INFORMATION:
      NAME: DeConti, Giulio A. Jr.
      REGISTRATION NUMBER: 31,503
      REFERENCE/DOCKET NUMBER: BBI-009C6CNDV
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (617)227-7400
      TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 7:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 520 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
   MOLECULE TYPE: DNA
   SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-489-777A-7

Query Match     5.4%;  Score 311.8;  DB 3;  Length 520;
Best Local Similarity   99.4%;  Pred. No. 1.7e-82;
Matches 313;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy  4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC  4336
         || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC   122

Qy  4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG  4396
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG   182

Qy  4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGT  4456
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGT   242

Qy  4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG  4516
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG   302

Qy  4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT  4576
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT   362

Qy  4577 CGAGCTCGGTACCAG  4591
         |||||||||||
Db   363 CGAGCTCGGTACCCG   377


RESULT 14
PCT-US95-08179-7
; Sequence 7, Application PC/TUS9508179
; GENERAL INFORMATION:
   APPLICANT:
   TITLE OF INVENTION:  Tetracycline-Regulated Transcriptional
   TITLE OF INVENTION:  Modulators
   NUMBER OF SEQUENCES:  28
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: LAHIVE & COCKFIELD
      STREET:  60 State Street, suite 510
      CITY:  Boston
      STATE:  Massachusetts
      COUNTRY:  USA
      ZIP:  02109-1875
   COMPUTER READABLE FORM:
      MEDIUM TYPE:  Floppy disk
      COMPUTER:  IBM PC compatible
      OPERATING SYSTEM:  PC-DOS/MS-DOS
      SOFTWARE:  ASCII Text
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER:  PCT/US95/08179
      FILING DATE:
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  To be assigned
      FILING DATE:  07-JUN-1995
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
      APPLICATION NUMBER:  US 08/383,754
      FILING DATE:  03-FEB-1995
      CLASSIFICATION:
   PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-94
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009C2PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US95-08179-7

Query Match   5.4%;   Score 311.8; DB 5; Length 520;
Best Local Similarity 99.4%;   Pred. No. 1.7e-82;
Matches 313;   Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
           ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db   63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy   4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 4576
           ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCAG 4591
           ||||||||| |
Db   363 CGAGCTCGGTACCCG 377

RESULT 15
US-08-076-726-13/c
; Sequence 13, Application US/08076726
; Patent No. 5464758
; GENERAL INFORMATION:
; APPLICANT: Gossen, Manfred
; APPLICANT: Bujard, Hermann
; TITLE OF INVENTION: Tight Control of Gene Expression in
; TITLE OF INVENTION: Eucaryotic Cells by Tetracycline-responsive Promoters
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: District of Columbia
; COUNTRY: United States of America
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,726
FILING DATE: 14-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2490001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
US-08-076-726-13

Query Match   5.4%;   Score 310.4; DB 1; Length 450;
Best Local Similarity 99.7%;   Pred. No. 4.1e-82;
Matches 311;   Conservative   0; Mismatches   0; Indels   1; Gaps   0;

Qy   4280 ACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCC 4339
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   325 ACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCC 266

Qy   4340 CTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAA 4399
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   265 CTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAA 206

Qy   4400 AGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTAC 4459
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   205 AGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTAC 146

Qy   4460 CACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATA 4519
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   145 CACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATA 86

Qy   4520 GAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGA 4579
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   85 GAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGA 26

Qy   4580 GCTCGGTACCAG 4591
           |||||||||||
Db   25 GCTCGGTACCCG 14

Search completed: April 25, 2005, 14:46:02
Job time : 643.574 secs

THIS PAGE BLANK (USPTO)

Montanari, D.
10/613728
Seq IDs (# 5-8

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 25, 2005, 00:13:02 ; Search time 16284 Seconds
                                           (without alignments)
                                           17065.210 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 ggatcctgcaaggtcacaca.........ccataagagttggagtcgac 5735

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   GenEmbl:*
                 1: gb_ba:*
                 2: gb_htg:*
                 3: gb_in:*
                 4: gb_om:*
                 5: gb_ov:*
                 6: gb_pat:*
                 7: gb_ph:*
                 8: gb_pl:*
                 9: gb_pr:*
                10: gb_ro:*
                11: gb_sts:*
                12: gb_sy:*
                13: gb_un:*
                14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5045	88.0	5443	6	BD140880	BD140880 An animal
2	5045	88.0	5443	10	MMU71441	U71441 Mus musculu
3	4804.8	83.8	235759	2	AC099577	AC099577 Mus muscu
4	2343	40.9	185702	2	AC130940	AC130940 Rattus no
c 5	2343	40.5	210784	2	AC119293	AC119293 Rattus no
6	2320.2	40.5	5487	10	AY191158	AY191158 Rattus no
c 7	1655.4	28.9	263901	2	AC115371	AC115371 Rattus no
8	1444.8	25.2	32415	10	HAMSHCA	L15351 Mesocricetu
9	1279.8	22.3	1679	6	AR054006	AR054006 Sequence
10	1279.8	22.3	1679	6	AR146182	AR146182 Sequence
11	1279.8	22.3	1679	6	AX546496	AX546496 Sequence
12	1279.8	22.3	2594	10	MUSCMHCG	M62404 Mus musculu
c 13	558.2	9.7	157910	9	CNS01DTR	AL132855 Human chr
c 14	558.2	9.7	196292	9	CNS0000B	AL049829 Human chr
15	494.8	8.6	31462	9	HSCAMHCA	Z20656 Homo sapien
16	462.6	8.1	516	11	BV093112	BV093112 RPAMMSEQ0
c 17	462.4	8.0	542	11	BV163870	BV163870 RPAMMSEQ0
c 18	438.4	7.6	452	11	BV093084	BV093084 RPAMMSEQ0
19	424.4	7.4	641	10	RATMYHAB1	K01463 Rat cardiac

20	407.6	7.1	6020	4	AF192305	AF192305 Oryctolag
c 21	396.4	6.9	541	11	BV159459	BV159459 RPAMMSEQ0
c 22	395.4	6.9	538	11	BV093111	BV093111 RPAMMSEQ0
c 23	378	6.6	580	11	BV159460	BV159460 RPAMMSEQ0
24	377	6.6	581	11	BV093113	BV093113 RPAMMSEQ0
25	343.6	6.0	2450	6	BD096863	BD096863 Agent for
26	312.6	5.5	10728	6	AR562497	AR562497 Sequence
27	311.8	5.4	520	6	AR020198	AR020198 Sequence
28	311.8	5.4	520	6	AR032150	AR032150 Sequence
29	311.8	5.4	520	6	AR043815	AR043815 Sequence
30	311.8	5.4	520	6	AR072127	AR072127 Sequence
31	311.8	5.4	520	6	AR095981	AR095981 Sequence
32	311.8	5.4	520	6	AR136641	AR136641 Sequence
33	311.8	5.4	520	6	AR157176	AR157176 Sequence
34	311.8	5.4	520	6	AR164125	AR164125 Sequence
35	311.8	5.4	520	6	I32811	I32811 Sequence 7
36	311.8	5.4	520	6	I59625	I59625 Sequence 7
37	311.8	5.4	520	6	AX127248	AX127248 Sequence v
38	310.8	5.4	5961	12	AF326320	AF326320 Shuttle v
c 39	310.4	5.4	450	6	AR020200	AR020200 Sequence
c 40	310.4	5.4	450	6	AR029415	AR029415 Sequence
41	310.4	5.4	450	6	AR032152	AR032152 Sequence
c 42	310.4	5.4	450	6	AR043817	AR043817 Sequence
c 43	310.4	5.4	450	6	AR072129	AR072129 Sequence
c 44	310.4	5.4	450	6	AR095983	AR095983 Sequence
c 45	310.4	5.4	450	6	AR136643	AR136643 Sequence

ALIGNMENTS

RESULT 1
BD140880
LOCUS   BD140880   5443 bp   DNA   linear   PAT 18-SEP-2002
DEFINITION   An animal into which p300 gene is introduced.
ACCESSION   BD140880
VERSION   BD140880.1 GI:23235825
KEYWORDS   WO 0205633-A/2.
SOURCE   Mus musculus (house mouse)
  ORGANISM   Mus musculus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 5443)
  AUTHORS   Hasegawa,K., Kawase,Y. and Suzuki,H.
  TITLE   An animal into which p300 gene is introduced
  JOURNAL   Patent: WO 0205633-A 2 24-JAN-2002;
             KOJI HASEGAWA,CHUGAI PHARMACEUTICAL CO LTD,YOSUKE KAWASE, HIROSHI
             SUZUKI
COMMENT   OS   Mus musculus (mouse)
             PN   WO 0205633-A/2
             PD   24-JAN-2002
             PF   13-JUL-2001 WO 2001JP006086
             PR   14-JUL-2000 JP 00P 215143
             PI   KOJI HASEGAWA,YOSUKE KAWASE,HIROSHI SUZUKI
             PC   A01K67/027,C12N15/09,C12N15/12,C12Q1/68,A61K45/00,A61P9/04, PC
                  G01N33/50,
             PC   G01N33/15
             CC   An animal into which p300 gene is introduced
             FH   Key   Location/Qualifiers
             FT   promoter   (1)..(5443).
FEATURES   Location/Qualifiers
     source   1..5443
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
ORIGIN

Query Match   88.0%; Score 5045; DB 6; Length 5443;
Best Local Similarity   93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

Qy.   1 GGATCCTGCAAGGTCACACATGCACACAAGGGTCTCCACCCACCAGGTGCCCAGTTCAT 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db      1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCCACCAGGTCTCCACCCAGTCTCAATTCAGT 60

Qy     61 TTCCATGCCTTGTTCTCACAAATGCTGGCCTCCCCAGAGCTAATTGGACTTTGTTTTTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCCATGCCTTGTTCTCACAAATGCTGGCCTCCCCAGAGCTAATTGGACTTTGTTTTTAT 120

Qy    121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTGCTACCCAGCTCTAAGGTGCCCGTGA 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTGCTACCCAGCTCTAAGGTGCCCGTGA 180

Qy    181 AGCCCTAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCCCTAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 240

Qy    241 TCCTTAAAGCCAAAATCCTGCCCTAGACTCTTCTTCTGACCTCGGTCCCTGGGCTCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCTTAAAGCCAAAATCCTGCCCTAGACTCTTCTTCTGACCTCGGTCCCTGGGCTCT 300

Qy    301 AGGGTGGGGAGGTGGGGCTTGGAAGAAGAAGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGGGTGGGGAGGTGGGGCTTGGAAGAAGAAGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360

Qy    361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCTCCAGCC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCTCCAGCC 420

Qy    421 AAACATAGCAAGAAGTGATACCTCCTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAACATAGCAAGAAGTGATACCTCCTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT 480

Qy    481 GAAACAGGATTTAGAGAAGCCTCTGAACTCACCTGAAGCTCTCACCACCAAGCA 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAAACAGGATTTAGAGAAGCCTCTGAACTCACCTGAAGCTCTCACCACCAAGCA 540

Qy    541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAGCTGGGCCAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAGCTGGGCCAC 600

Qy    601 AGAAGTCCTGGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGGTATGACCCCC 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGAAGTCCTGGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGGTATGACCCCC 660

Qy    661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCATCCCAGTTCTTCACGTCTCCTCAAGCA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCATCCCAGTTCTTCACGTCTCCTCAAGCA 720

Qy    721 GGATGAGAGATGTAGTCATGTGGCATTCCAAACACAGTCTCCACAGTGTCCCCTGCCCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGATGAGAGATGTAGTCATGTGGCATTCCAAACACAGTCTCCACAGTGTCCCCTGCCCC 780

Qy    781 TTCCCACTTAGCCAGGAGGACAGTGACTTAGCCTATCTTCTTCCTCACGTCTCCCATCCTCCCCAG 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTCCCACTTAGCCAGGAGGACAGTGACTTAGCCTATCTTCTTCCTCACGTCTCCCATCCTCCCCAG 840

Qy    841 GACACACCCCCTGGTCTGCGAGTATTCATTTCATTTCTTCCTTCACGTCGACTTCCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACACACCCCCTGGTCTGCGAGTATTCATTTCATTTCTTCCTTCACGTCGACTTCCAT 900

Qy    901 TTGCAAGGCTTTTGACCTCTGCAGCTGCTGGGCTGCTGGAAGTTTGGCCCTAGGTGTGGCAA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGCAAGGCTTTTGACCTCTGCAGCTGCTGGGCTGCTGGAAGTTTGGCCCTAGGTGTGGCAA 960

Qy    961 GCCATCTCAAGAGAAAGCAAGCAGACAACACAGCAGGGGGACCAGATTTTGGAAGGATCAGGAACTAAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCCATCTCAAGAGAAAGCAAGCAGACAACACAGCAGGGGGACCAGATTTTGGAAGGATCAGGAACTAAA 1020

Qy   1021 TCACTGGCCGGGGCTGGGGGTAGAAAAAAAGAGTGAGTCCGCTCAGCTAAGGCCAAGC 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCACTGGCCGGGGCTGGGGGTAGAAAAAAAGAGTGAGTCCGCTCAGCTAAGGCCAAGC 1080

Qy   1081 TAGTCCCCGAGATATACTCTGCCACAGTCTGGGCGTGCTGCTCGGGGTAGCTTTGGGGTC 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAGTCCCCGAGATATACTCTGCCACAGTCTGGGCGTGCTGCTCGGGGTAGCTTTGGGGTC 1140
```

```
Qy   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACCTACAGAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACCTACAGAC 1200

Qy   1201 ACACTCGTCGTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTTCGTTCTGACCACTGAGCTAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACACTCGTCGTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTTCGTTCTGACCACTGAGCTAG 1260

Qy   1261 GCAACCAGAGCATGGCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCAACCAGAGCATGGCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320

Qy   1321 CAGGGGAGGGAGACAGAGCAGGAAATAAGGAAGAAGGAAGGCCAGTCAATCAGA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CAGGGGAGGGAGACAGAGCAGGAAATAAGGAAGAAGGAAGGCCAGTCAATCAGA 1380

Qy   1381 TGCAGTCAGAAGAGATGGGAAGCCAACACACAGCTTGACGCAGAAACAGAAAAGGGAG 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCAGTCAGAAGAGATGGGAAGCCAACACACAGCTTGACGCAGAAACAGAAAAGGGAG 1440

Qy   1441 AGATTCTGGGCATAAGGAGGCCACAGAAGAAGAGAGCCCAGCCCCCCAAGTCTCTCTTT 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AGATTCTGGGCATAAGGAGGCCACAGAAGAAGAGAGCCCAGCCCCCCAAGTCTCTCTTT 1500

Qy   1501 ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560

Qy   1561 AGCGAAGAGACCCGTAGGGAGGATTCTGGGGGGGGGATTCTGGGGGGGGGGGGGAGAAGTGCCAGG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGCGAAGAGACCCGTAGGGAGGATTCTGGGGGGGGGATTCTGGGGGGGGGGGGGAGAAGTGCCAGG 1620

Qy   1621 GCAACCTAAGAGCCAGAGCCTAAAAGAGCAAGACAAGTAAAGGTGCTTCAAAGGTGGCCAGG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCAACCTAAGAGCCAGAGCCTAAAAGAGCAAGACAAGTAAAGGTGCTTCAAAGGTGGCCAGG 1680

Qy   1681 CTGTGCACACAGAGGGTCGAGGACTGGTGTGTAGAGCCTCAAGATAAGGATGATGCTCAGA 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CTGTGCACACAGAGGGTCGAGGACTGGTGTGTAGAGCCTCAAGATAAGGATGATGCTCAGA 1740

Qy   1741 ATGGGCGGGGGGGGGGGGATTCTGGGGGGATTCTGGGGGGGGGGAGAGGAGCCTGAACAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATGGGCGGGGGGGGGGGGATTCTGGGGGGATTCTGGGGGGGGGGAGAGGAGCCTGAACAG 1800

Qy   1801 AGAATCTGGAAGCGCTGGAAACGTGGAAACGTGGAAACAGCAGAGACCCAGGCTACCTTTAGATG 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGAATCTGGAAGCGCTGGAAACGTGGAAACGTGGAAACAGCAGAGACCCAGGCTACCTTTAGATG 1860

Qy   1861 TAAATCATGAAAGACAGGGAAGGGAAGCAGGGGAGCCAGGTAGCAAAGGAATGGTTCCTTCTCAC 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TAAATCATGAAAGACAGGGAAGGGAAGCAGGGGAGCCAGGTAGCAAAGGAATGGTTCCTTCTCAC 1920

Qy   1921 CATGGAAGCAAGGACAAGCCAGGTTGAGCCGCTGCTGCTGAAATCAGCCTGCTGAAGGCAGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CATGGAAGCAAGGACAAGCCAGGTTGAGCCGCTGCTGCTGAAATCAGCCTGCTGAAGGCAGAG 1980

Qy   1981 CCCTGGTATGAGCACCAGAACAGCAGACAGAGGGTAGGGTTAATGTCGAGACAGGGAACAGAAG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CCCTGGTATGAGCACCAGAACAGCAGACAGAGGGTAGGGTTAATGTCGAGACAGGGAACAGAAG 2040

Qy   2041 GTAGACACAGGAACAGACAGAGACGGGGAGCCAGGGGAGCCAGGTAACAAAGGAATGGTCCTTCTCAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GTAGACACAGGAACAGACAGAGACGGGGAGCCAGGGGAGCCAGGTAACAAAGGAATGGTCCTTCTCAC 2100

Qy   2101 CTGTTGGCCAGAGCTGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGGC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGTTGGCCAGAGCTGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGGC 2160

Qy   2161 TGGCTTGGGAGGCAGCTGGAAAGTATGTGAGACGCAAGGGGAGCAAGGGGAGCCTAGGA 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TGGCTTGGGAGGCAGCTGGAAAGTATGTGAGACGCAAGGGGAGCAAGGGGAGCCTAGGA 2220
```

```
Qy  3301  CTGGCCTTCCCCTTCCCACCTCCCACCTCCCCATCCCAGCGAGTGGAGGGTTGCAGAGGGTAAAAA  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CTGGCCTTCCCCTTCCCACCTCCCACCTCCCCATCCCAGCGAGTGGAGGGTTGCAGAGGGTAAAAA  3360

Qy  3361  CCTACATGTCCAAACATCATGTGCACGATATATGTGCGATATGTGACATCAGTCATGGACATGT  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361  CCTACATGTCCAAACATCATGTGCACGATATATGTGCGATATGTGACATCAGTCATGGACATGT  3420

Qy  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTCTG  3480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTCTG  3480

Qy  3481  ACTGAAAACGGCATGGCTGTGCAGCTGTTCAGTTCTCGTGCGTGAGCTTACCAGACTGCA  3540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  ACTGAAAACGGCATGGCTGTGCAGCTGTTCAGTTCTCGTGCGTGAGCTTACCAGACTGCA  3540

Qy  3541  GGTTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCATGGTTTAAAGAGAT  3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GGTTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCATGGTTTAAAGAGAT  3600

Qy  3601  TGGATGATGGCCTGCATCTCAAGGACCACCTGTCTTTGGAGGACGATGTGGCAACAGAGAC  3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  TGGATGATGGCCTGCATCTCAAGGACCACCTGTCTTTGGAGGACGATGTGGCAACAGAGAC  3660

Qy  3661  CTAAGCTAAGGTAGCAAGGTCTTTGGAGGTCAGAGTCTCCTTGGAGAGTAGGAGGCTCCCCTCA  3720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  CTAAGCTAAGGTAGCAAGGTCTTTGGAGGTCAGAGTCTCCTTGGAGAGTAGGAGGCTCCCCTCA  3720

Qy  3721  TACAGACAGTATCTGTACAGAGTAAGGACAGGGGTGTAGAATTCTCTTACTA  3780
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  TACAGACAGTATCTGTACAGAGTAAGGACAGGGGTGTAGAATTCTCTTACTA  3780

Qy  3781  TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGATGCTCCCTAGACATCATCATGACTT  3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGATGCTCCCTAGACATCATGACTT  3840

Qy  3841  TGTCTCTGGGAGCCAGCACTGTGAACTTCAGGTCGAGAGTAGGAGGCTCCCTCA  3900
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTCTCTGGGAGCCAGCACTGTGAACTTCAGGTCGAGAGTAGGAGGCTCCCTCA  3900

Qy  3901  GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAAGGGAGACCTGGGATGGGAGC  3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAAGGGAGACCTGGGATGGGAGC  3960

Qy  3961  TTGTGTGTTGGAGGGGTAGGGGAAGGCAGAGATTATTAAGCCTGGACCCTTACCCAG  4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  TTGTGTGTTGGAGGGGTAGGGGAAGGCAGAGATTATTAAGCCTGGACCCTTACCCAG  4020

Qy  4021  TTGTTCAACTCACCCTTCAGATAAAAATAACTGAGGTAAGGGCCTGGGTAGGGACAGTG  4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTGGGTAGGGACAGTG  4080

Qy  4081  GTGTGAGACCGCTCCTGTCTCCTCTGCTTGGGGTGTGCTTGGGGGAGGAGGAATGT  4140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTGTGAGACCGCTCCTGTCTCCTCTGCTTGGGGTGTGCTTGGGGGAGGAGGAATGT  4140

Qy  4141  GCCCAAGGACTAAAAAAGGCCATGGACCCGAGCGAGGGCGACCCTTTCATGG  4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  GCCCAAGGACTAAAAAAGGCCATGGACCCGAGCGAGGGCGACCCTTTCATGG  4200

Qy  4201  GCAAACCTTGGGGCCCGTAGTGATCCGAGTTTACCGATCAAGGGATCAAAGGAGGA  4260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  GCAAACCTTGGGGCCCGTCGTCCTCCTGTCACCTCCGATCAAGGGATCAAAGGAGGA  4260

Qy  4261  TTCTAACGGACAGGAGGACTCGAGTTTACCACTCCCTATCAGTGATGAGAAAGTGA  4320
          ||||||||||||||||||||
Db  4261  GGAGCCAGGAGGACGAGGAGTGGGAGGGAG---------------------------  4292

Qy  4321  AGTCGGAGTTTACCACTCCCTATCAGTGATGAGAAAGTCGAGTTTACCACTC  4380
Db  4292  --------------------------------------------------  4292

Qy  4381  CCTATCAGTGATGATAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAA  4440
```

```
Qy  2221  AAGGAAGAAGAGAGGGCAAAACCAGGCCACACAAGAGGGCAGAGCCCAGAACTGAGTTAACTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  AAGGAAGAAGAGAGGGCAAAACCAGGCCACACAAGAGGGCAGAGCCCAGAACTGAGTTAACTC  2280

Qy  2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA  2340

Qy  2341  GCCCCCACTACCCATACCAAGTTTGGCTGGCATTCTAGGTTCCCTGAGGACAGAG  2400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GCCCCCACTACCCATACCAAGTTTGGCTGGCATTCTAGGTTCCCTGAGGACAGAG  2400

Qy  2401  CCTGGCCTTTGTCTCTTGACCTGACCCCAAGCAAGTCTCAGTACCTTTGAAT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CCTGGCCTTTGTCTCTTGACCTGACCCCAAGCAAGTCTCACGTACCTTTATCAT  2460

Qy  2461  GCCCTCAAGAGCTTGAGAACCAGCAGTGACATATTAGGCTACAACCCTGGAGCT  2520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  GCCCTCAAGAGCTTGAGAACCAGCAGTGACATATTAGGCTACAACCCTGGAGCT  2520

Qy  2521  TGCACAGGAGGCTCAGTGACCTTGCCAGGACACAGCTGCAGACAGTGGCCTTTATCC  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  TGCACACAGGAGGCTCAGTGACCTTGCCAGGACACAGCTGCAGACAGTGGCCTTTATCC  2580

Qy  2581  CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGAATGGAATCAGGTGCAAATCAGGTC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGAATGGAATCAGGTGCAAATCAGGTC  2640

Qy  2641  CCTTCAAGATATACTGCATGCAGACCTAAGACCCCTGGAGAGAGGGTATGCTCCTGCCC  2700
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  CCTTCAAGATATACTGCATGCAGACCTAAGACCCCTGGAGAGAGGGTATGCTCCTGCCC  2700

Qy  2701  CCACCCACCATAAGGGGAGTGAACTATCCTAGGGGCTGGCACCTTGGGGGACACCAC  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  CCACCCACCATAAGGGGAGTGAACTATCCTAGGGGCTGGCACCTTGGGGGACACCAC  2760

Qy  2761  ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCCGTCGTCTGCGCCACCTCCACAC  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCCGTCGTCTGCGCCACCTCCACAC  2820

Qy  2821  TCTAGAGCTATATTGAGAGGTGACAGTAGAGGTGGGAGCTCGTGAGCAGGAGAGTGT  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TCTAGAGCTATATTGAGAGGTGACAGTAGAGGTGGGAGCTCGTGAGCAGGAGAGTGT  2880

Qy  2881  TCCTGGGGTGTGAGGGGTAGGGGAAAGCCAGAGGGAGTCTCGGCTTTGTCTCCTGAA  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  TCCTGGGGTGTGAGGGGTAGGGGAAAGCCAGAGGGAGTCTCGGCTTTGTCTCCTGAA  2940

Qy  2941  CACAATGTCTACTTAGTTATAACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC  3000
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  CACAATGTCTACTTAGTTATAACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC  3000

Qy  3001  TGAAAAGACAGCAGCCCTGGAGGACAGGGGTTGTCTCTGGCTTGATGGTG  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TGAAAAGACAGCAGCCCTGGAGGACAGGGGTTGTCTCTGGCTTGATGGTG  3060

Qy  3061  CCACAAAGGAGGGCCATGAGTGTGAGTATAGGCTGAGGCGTTAGAAGGCACTTG  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  CCACAAAGGAGGGCCATGAGTGTGAGTATAGGCTGAGGCGTTAGAAGGCACTTG  3120

Qy  3121  GGAGGGGGTCAGTCTGCAGAGCCCTATCCATGAATCTGAGGCTGGGCCAACTGGTG  3180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  GGAGGGGGTCAGTCTGCAGAGCCCTATCCATGAATCTGAGGCTGGGCCAACTGGTG  3180

Qy  3181  TAAATCTCTGGGGCCTGCCAGCCCCTATACCCTTTCTCCCCCAGATTAACACCT  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  TAAATCTCTGGGGCCTGCCAGCCCCTATACCCTTTCTCCCCCAGATTAACACCT  3240

Qy  3241  GGCGAAGGGAGCAACCCCCACCCCCAGCCCCAGATTAACACCT  3300
          |||||||||||||||||||||||||||||||||||||||||||
Db  3241  GGCGAAGGGAGCAACCCCCACCCCCAGCCCCAGATTAACACCT  3300
```

Left column (continuation of alignment):

```
Db  4292 ------------------------------------------------------------ 4291
Qy  4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA 4500
Db  4292 ------------------------------------------------------------ 4291
Qy  4501 CCACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGAT       4560
Db  4292 ------------------------------------------------------------ 4291
Qy  4561 AGAGAAAGTGAAAGTCGAGTCGGTACCAGCAGAGACTCCAAATTTAGGCAGCAGGCA     4620
Db  4292 --------GGTCCCAGCAGCAGGAGCACTGAGAGCTCCAAATTTAGGCAGCAGGCA       4328
         |||||||||||||||||||||||
Qy  4621 TATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGCTGTCAGAGATTTCTCCAACCCA  4680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4329 TATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGCTGTCAGAGATTTCTCCAACCCA  4388
Qy  4681 GGTAAGAGGGAGTTTCGGGTGGGGGGGTCTTCACCCACCACCAGACCTCTCCCCACCTAGAA 4740
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4389 GGTAAGAGGGAGTTTCGGGTGGGGGGGTCTTCACCCACCACCAGACCTCTCCCCACCTAGAA 4448
Qy  4741 GAAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCTGACAGGGTGGCCTT  4800
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4449 GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCTGACAGGGTGGCCTT  4508
Qy  4801 CCACCAGCCTGGGAAGTTCTCAGTGCAGCAGGAGGTTCCACAAGAAAACTCGGATGCCCT  4860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4509 CCACCAGCCTGGGAAGTTCTCAGTGCAGCAGGAGGTTCCACAAGAAAACTCGGATGCCCT  4568
Qy  4861 TCCCTTACGCTGTCTTCTCCATCTCCCGGGATGTCCTCCCGTCTTGGTTTATC         4920
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4569 TCCCTTACGCTGTCTTCTCCATCTCCCGGGATGTCCTCCCGTCTTGGTTTATC         4628
Qy  4921 TTGGCTCTTCGTCTTCAGCAAGATTGCCCTGTGCTGTCGTCCACTCCATCTTTCTCTACTGT 4980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4629 TTGGCTCTTCGTCTTCAGCAAGATTGCCCTGTGCTGTCGTCCACTCCATCTTTCTCTACTGT 4688
Qy  4981 CTCCGTGCCTGCCTTGCCTTGCGGTGTCCTTCGTTCCACCCATTCTCACTTCACC       5040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4689 CTCCGTGCCTGCCTTGCCTTGCGGTGTCCTTCGTTCCACCCATTCTCACTTCACC       4748
Qy  5041 TTTTCTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC         5100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4749 TTTTCTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC         4808
Qy  5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC          5160
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4809 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC          4868
Qy  5161 CTTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATGTAAACAATCTTCCA  5220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4869 CTTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCCTTATGTAAACAATCTTCCA 4928
Qy  5221 GTGAGCCCAGCTTCAGTGCGTCGGTGGGTGCTCTCTTCACCCCTGGCTTGTC          5280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4929 GTGAGCCCAGCTTCAGTGCGTCGGTGGGTGCTCTCTTCACCCCTGGCTTGTC          4988
Qy  5281 CTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCCTCCT  5340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4989 CTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCCTTCCT 5048
Qy  5341 GCCTGTTCCTCTCTGTCCAGCTGCGCCACTGCTGCCTCGTTCCAGCTGTCGGTCCAC     5400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5049 GCCTGTTCCTCTCTGTCCAGCTGCGCCACTGCTGCCTCGTTCCAGCTGTCGGTCCAC     5108
Qy  5401 ATTCTTCAGGATTCTCCCGGAAGTTCCCGGAAGTTAACCAGGTGAAAATGTTTCCCCTGTAGACAGAGA 5460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5109 ATTCTTCAGGATTCTCCCGGAAGTTCCCGGAAGTTAACCAGGTGAAAATGTTTCCCCTGTAGACAGAGA 5168
Qy  5461 TCACGATTCTCCCGGAAGTCAGGCTTCAGCCTTCCAGCCCTCTCTTTCTGTTTCTGCCCAGCTGCCCGGCA 5520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5169 TCACGATTCTCCCGGAAGTCAGGCTTCAGCCTTCCAGCCCTCTCTTTCTGTTTCTGCCCAGCTGCCCGGCA 5228
```

Right column (end of alignment):

```
Qy  5521 CTCTTAGCAAACCTCAGGCACCCTTACCCCACCATAGACCTCTGACAGAGAAGCAGGCACT 5580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5229 CTCTTAGCAAACCTCAGGCACCCTTACCCCACCATAGACCTCTGACAGAGAAGCAGGCACT 5288
Qy  5581 TTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGAGGCAGGGAAGTGGT  5640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5289 TTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGAGGCAGGGAAGTGGT  5348
Qy  5641 GGTGTAGGAAAGTCAGGACTTCACATAGAGAGCCTAGCCCACCAGAAATGACGACAGA    5700
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5349 GGTGTAGGAAAGTCAGGACTTCACATAGAGAGCCTAGCCCACCAGAAATGACGACAGA    5408
Qy  5701 TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC 5735
         ||||||||||||||||||||||||||||||||||
Db  5409 TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC 5443
```

RESULT 2
MMU71441
LOCUS   MMU71441   5443 bp   DNA   linear   ROD 21-OCT-1996
DEFINITION Mus musculus alpha myosin heavy chain gene, promoter region.
ACCESSION   U71441
VERSION   U71441.1 GI:1621436
KEYWORDS   .
SOURCE   Mus musculus (house mouse)
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 5443)
  AUTHORS   Gulick,J. and Robbins,J.
  TITLE   Mouse myosin heavy chain promoter
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 5443)
  AUTHORS   Gulick,J. and Robbins,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-SEP-1996) Molecular Card. Biol., Children's Hospital,
            3333 Burnet Avenue, Cincinnati, OH 45229-3039, USA
FEATURES   Location/Qualifiers
     source   1..5443
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     promoter   1..5443
                     /note="alpha myosin heavy chain; capable of driving high
                     levels of cardiac specific transcription in vitro or in
                     vivo"
ORIGIN

```
Query Match             88.0%;  Score 5045;  DB 10;  Length 5443;
Best Local Similarity   93.9%;  Pred No. 0;
Matches  5383;  Conservative    0;  Mismatches    0;  Indels  292;  Gaps    1;

Qy    1 GGATCCTGCAAGGTCACACAAGGGTCTCACCCACCACCAGGTGCCCTAGTCTCAATTTCAGT  60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GGATCCTGCAAGGTCACACAAGGGTCTCACCCACCACCAGGTGCCCTAGTCTCAATTTCAGT  60
Qy   61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCAGAGAGCTAATTTGGACTTTGTTTTTAT  120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCAGAGAGCTAATTTGGACTTTGTTTTTAT  120
Qy  121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTGCTCTAAGGTGCCCGTGA         180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTGCTCTAAGGTGCCCGTGA         180
Qy  181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTTGGAAGCAGAATAAAGCAATTT  240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTTGGAAGCAGAATAAAGCAATTT  240
Qy  241 TCCTTAAAGCCAAAATCCTGCCTCTAGACTCTTCTTCTCTGACCTCGGTCCGGTCGGGCTCT 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  241 TCCTTAAAGCCAAAATCCTGCCTCTAGACTCTTCTTCTCTGACCTCGGTCCGGTCGGGCTCT 300
```

```
Db 1381 TGCAGTCAGAGAGAGATGGGAGGCCAACACAGCTTGAGCAGAGGAAACACAGAAAAGGGAG 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 1441 AGATTCTGGGCATAAGGAGGAGCCACAGAGGAGCCCAGCCCCCCCAGTCTCCTCTTT 1500
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441 AGATTCTGGGCATAAGGAGGAGCCACAGAGGAGCCCAGCCCCCCCAGTCTCCTCTTT 1500

Qy 1501 ATACCCTCATCCGTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501 ATACCCTCATCCGTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560

Qy 1561 AGCGAAGAGACCCGTAGGAGGGATCACACTCGGATGAAGGGAGGATGTGTGGAGAAGTCCAGG 1620
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561 AGCGAAGAGACACCCGTAGGAGGGATCACACTCGGATGAAGGGAGGATGTGTGGAGAAGTCCAGG 1620

Qy 1621 GCAACCTAAGAGCCAGAGCCTAAAAGCAAGATAAAGTGCTTCAAAGGTGGCCAGG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621 GCAACCTAAGAGCCAGAGCCTAAAAGCAAGATAAAGTGCTTCAAAGGTGGCCAGG 1680

Qy 1681 CTGTGCACAGAGGGTCGAGGACTGTGGTGAGAGCCTCAAGATAAGGATGATGCTCAGA 1740
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681 CTGTGCACAGAGGGTCGAGGACTGTGGTGAGAGCCTCAAGATAAGGATGATGCTCAGA 1740

Qy 1741 ATGGGCGGGGGGGGGGGGGATTCTGGGGGGGGAGGAGGCTGAGAAGGAGGCTGAACAG 1800
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1741 ATGGGCGGGGGGGGGGGGGATTCTGGGGGGGGAGGAGGCTGAGAAGGAGGCTGAACAG 1800

Qy 1801 AGAATCTGGAAGCGCTGGAAACGATCACATAAAGGGAAGAACCCCAGGCTACCTTTAGATG 1860
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1801 AGAATCTGGAAGCGCTGGAAACGATCACATAAAGGGAAGAACCCCAGGCTACCTTTAGATG 1860

Qy 1861 TAAATCATGAAAGACAGGGAGAAGGAAGTGGAGAGTGAAGGAGACCCGGGCCAAGA 1920
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1861 TAAATCATGAAAGACAGGGAGAAGGAAGTGGAGAGTGAAGGAGACCCGGGCCAAGA 1920

Qy 1921 CATGGAAGCAAGGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGCAGAGG 1980
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921 CATGGAAGCAAGGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGCAGAGG 1980

Qy 1981 CCCTGTATGAGCACCAGAGAACAGCAGAGGTTAGGGTTAAATGTCGAGACAGGGAAG 2040
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981 CCCTGTATGAGCACCAGAGAACAGCAGAGGTTAGGGTTAAATGTCGAGACAGGGAAG 2040

Qy 2041 GTAGACACAGGAACAGCAGAGACAGGGGAGCCAGTAACAAAGGAATGGTCCTTCTCAC 2100
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041 GTAGACACAGGAACAGCAGAGACAGGGGAGCCAGTAACAAAGGAATGGTCCTTCTCAC 2100

Qy 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCGAGACTGCGAGGGC 2160
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCGAGACTGCAGGGGC 2160

Qy 2161 TGGCTTGGGAGGCAGCTGGAAAGAGTATGTGAGAAGACAAGGGGCCTAGGA 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161 TGGCTTGGGAGGCAGCTGGAAAGAGTATGTGAGAAGACAAGGGGCCTAGGA 2220

Qy 2221 AAGGAAGAAGAGGGCAAACCAGGCCACAAGAGGGCCAGCCCCAGAACTGAGTTAACTC 2280
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221 AAGGAAGAAGAGGGCAAACCAGGCCACAAGAGGGCCAGCCCCAGAACTGAGTTAACTC 2280

Qy 2281 CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATGA 2340
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATGA 2340

Qy 2341 GCAACCCACTACCCATACCAAGTTTGGCCTGAGTGGCATTCTAGGTTCCCTGAGGACAGAG 2400
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 GCAACCCACTACCCATACCAAGTTTGGCCTGAGTGGCATTCTAGGTTCCCTGAGGACAGAG 2400

Qy 2401 CCTGGCCTTTGTCTTCTTGGACCTGACCCAAGCTGACCCAATGTTCTCAGTACCTTTGAAT 2460
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 CCTGGCCTTTGTCTTCTTGGACCTGACCCAAGCTGACCCAATGTTCTCAGTACCTTTATCAT 2460

Qy 2461 GCCCTCAAGAGCTTGAGAACCAGGCAGTGACATATTAGGCCATGGGCTAACCCTGGAGCT 2520
```

```
Qy 301 AGGGTGGGGAGGTGGGGCTTGGAGAAGAAGAGGTCGGGGAGTGGGCAAAAGCCGATCCCTAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 AGGGTGGGGAGGTGGGGCTTGGAGAAGAAGAGGTCGGGGAGTGGGCAAAAGCCGATCCCTAG 360

Qy 361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTTCCAGCC 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTTCCAGCC 420

Qy 421 AAACATAGACAAGAAGTGATACCTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 AAACATAGACAAGAAGTGATACCTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT 480

Qy 481 GAAACAGGATTTAGAGAAGAGCCTCTGAACTCACCTGAACTGAAGCTCATCCACCAAGCA 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 GAAACATGGATTTAGAGAAGAGCCTCTGAACTCACCTGAACTGAAGCTCATCCACCAAGCA 540

Qy 541 AGCACCTAGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAGGTATGACCCCAC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 AGCACCTAGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAGGTATGACCCCAC 600

Qy 601 AGAAGTCCTGGGGTGTAGGAACTGACCCAGTGACTTTTCAGTCGGCAAAGTATCCCCC 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 601 AGAAGTCCTGGGGTGTAGGAACTGACCCAGTGACTTTTCAGTCGGCAAAGTATCCCCC 660

Qy 661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCCAGGCAGTCTCCAGGCAGTCTCTTAAGAGGACATG 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCCAGGCAGTCTCCAGGCAGTCTCTTAAGAGGACATG 720

Qy 721 GGATGAGAGATGTAGTCATGTGGCATTCCAAACACAGCTATCCACAGTGTCCCCTGCCCC 780
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 721 GGATGAGAGATGTAGTCATGTGGCATTCCAAACACAGCTATCCACAGTGTCCCCTGCCCC 780

Qy 781 TTTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTTCTTCCTCCCCATCCTCCCAG 840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 781 TTTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTTCTTCCTCCCCATCCTCCCAG 840

Qy 841 GACACACCCCTGGTCTCGCAGTATTCATTTCTTCCACGTCCCCTCTGTGACTTCCAT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 841 GACACACCCCTGGTCTCGCAGTATTCATTTCTTCCACGTCCCCTCTGTGACTTCCAT 900

Qy 901 TTGCAAGGCTTTGACCTCTGCAGCTGCTGAAGATGAGTTTGGCCCTAGGTGTGGCAA 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 901 TTGCAAGGCTTTGACCTCTGCAGCTGCTGAAGATGAGTTTGGCCCTAGGTGTGGCAA 960

Qy 961 GCCATCTCAAGAGAAAGCAGACAGACACAGGGGACCAGATTTTGGAAGGATCAGGAACTAAA 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 961 GCCATCTCAAGAGAAAGCAGACAGACACAGGGGACCAGATTTTGGAAGGATCAGGAACTAAA 1020

Qy 1021 TCACTGGCCGGCCTGGGGGTAGAAAAAAGAGTGAGTGAGTCCGCTCCAGTCAAGCCAAGC 1080
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021 TCACTGGCCGGCCTGGGGGTAGAAAAAAGAGTGAGTGAGTCCGCTCCAGTCAAGCCAAGC 1080

Qy 1081 TAGTCCCCAGAGATACTCTGCCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081 TAGTCCCCAGAGATACTCTGCCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC 1140

Qy 1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACTACAGAC 1200
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACTACAGAC 1200

Qy 1201 ACACTCGTGTGTGCCAGACTCCCTGTTCAACAGCCCTGAGTTCTGAGGATGAGAGTTGGTTACCAATAGCAAAAACAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201 ACACTCGTGTGTGCCAGACTCCCTGTTCAACAGCCCTGAGTTCTGAGGATGAGAGTTGGTTACCAATAGCAAAAACAG 1260

Qy 1261 GCAACCAGAGCATGGGCGCCCTGTCGTGAGGTAGAAATAGAGTTGGTTACCAATAGCAAAAACAG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261 GCAACCAGAGCATGGGCGCCCTGTCGTGAGGTAGAAATAGAGTTGGTTACCAATAGCAAAAACAG 1320

Qy 1321 CAGGGGAGGGAGAACAGAGAACGAAATAAGGAAGAAGGCCAGTCAATCAGA 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321 CAGGGGAGGGAGAACAGAGAACGAAATAAGGAAGAAGGCCAGTCAATCAGA 1380

Qy 1381 TGCAGTCAGAGAGAGATGGGAGGCCAACACAGCTTGAGCAGAGGAAACACAGAAAAGGGAG 1440
```

```
Db  2461 GCCCTCAGAGCTTGAGACCAGCAGTGACATATTAGGCCATGGGCTAACCCTGGAGCT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2521 TGCACACAGAGGAGCTCAGGTGACCTGCAGACAGCTGCAGACAGCTGGCCTTTATCC 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 TGCACACAGAGGAGCTCAGGTGACCTGCAGACAGCTGCAGACAGCTGGCCTTTATCC 2580

Qy  2581 CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGAATGCAAGGTTGAATCAGTC 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGAATGCAAGGTTGAATCAGTC 2640

Qy  2641 CCTTCAAGAATACTGCATGCAAGACCCTGAGAGAGGGTATGCTCCTGCCC 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641 CCTTCAAGAATACTGCATGCAAGACCCTGAGAGAGGGTATGCTCCTGCCC 2700

Qy  2701 CCACCCCATAGGGGAGTGAACTATCTCAGGGGCTGGCGACCTTGGGGACCACCAC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701 CCACCCCATAGGGGAGTGAACTATCTCAGGGGCTGGCGACCTTGGGGACCACCAC 2760

Qy  2761 ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCCGTGTCTGCCCACCTCCACAC 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761 ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCCGTGTCTGCCCACCTCCACAC 2820

Qy  2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGTAGCAGGGAGAGTGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGTAGCAGGGAGAGTGT 2880

Qy  2881 TCCTGGGTGTGAGGGTGTAGGGGGAAAGCCAGCAGAGCGAGTCTGGCTTGTCTCCTGAA 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881 TCCTGGGTGTGAGGGTGTAGGGGGAAAGCCAGCAGAGCGAGTCTGGCTTGTCTCCTGAA 2940

Qy  2941 CACAATGTCTACTTAGTTATAAACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC 3000
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941 CACAATGTCTACTTAGTTATAAACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC 3000

Qy  3001 TGAAAAGACAGCAGCCCTGGAGACACAGGGGTGTCTCTGAGCCTTGGGTGCTTGATGGTG 3060
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001 TGAAAAGACAGCAGCCCTGGAGACACAGGGGTGTCTCTGAGCCTTGGGTGCTTGATGGTG 3060

Qy  3061 CCCAAAGGAGGGGCATGAGTGTGAGTATAAGGCCCCAGGAGGTTAGAGAAGGGCACTTG 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061 CCACAAAGGAGGGGCATGAGTGTGAGTATAAGGCCCCAGGAGGTTAGAGAAGGGCACTTG 3120

Qy  3121 GGAAGGGGTCAGTCTGCAGAGCCCCTATCGAGAGTGCATGGAATCTGAGCCTGGGGCCAACTGGTG 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121 GGAAGGGGTCAGTCTGCAGAGCCCCTATCGAGAGTGCATGGAATCTGAGCCTGGGGCCAACTGGTG 3180

Qy  3181 TAAATCTCTGGGGCTGCCAGGCATTCAAAGCAGCACCTGCATCTCTGCAGCCTGGGGA 3240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181 TAAATCTCTGGGGCTGCCAGGCATTCAAAGCAGCACCTGCATCTCTGCAGCCTGGGGA 3240

Qy  3241 GGGCGAAGGGAGCAACCCCCCACTTATACCCCACTTATACCCTTTCTCCCAGGATTAAACACCT 3300
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241 GGGCGAAGGGAGCAACCCCCCACTTATACCCCACTTATACCCTTTCTCCCAGGATTAAACACCT 3300

Qy  3301 CTGGCCTTCCCCCTTCCCACCTCCCATCAGGAGTGGAGGGTTGAGAGGGAGGGTAAAAA 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301 CTGGCCTTCCCCCTTCCCACCTCCCATCAGGAGTGGAGGGTTGAGAGGGAGGGTAAAAA 3360

Qy  3361 CCTACATGTCCAAACATCATGGTTGCACGATATATGGATCAGTATGTGTAGAGGCCAAGAAA 3420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3361 CCTACATGTCCAAACATCATGGTTGCACGATATATGGATCAGTATGTGTAGAGGCAAGAAA 3420

Qy  3421 GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTCAGTTCTGTCATGTGTGTGTCTG 3480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421 GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTCAGTTCTGTCATGTGTGTGTCTG 3480

Qy  3481 ACTGAAAACGGGCATGGGCATGCTGTTCAGTTCGTGCGTGAGGTTACCAGACTGCA 3540
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481 ACTGAAAACGGGCATGGGCATGCTGTTCAGTTCGTGCGTGAGGTTACCAGACTGCA 3540

Qy  3541 GGTTTGTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541 GGTTTGTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
```

```
Db  3601 TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAAATGACACTCTATATGTGTCT 3660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3601 TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAAATGGACACTCTATATGTGTCT 3660

Qy  3661 CTAAGCTAAGGTAGTAGCAAGGTCTTTGGAGGACACCTGTCTTAGAGATGTGGCAACAGAGAC 3720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661 CTAAGCTAAGGTAGTAGCAAGGTCTTTGGAGGACACCTGTCTTAGAGATGTGGCAACAGAGAC 3720

Qy  3721 TACAGACAGTATCTGTACAGAGTAAGGAGGAGAGGGGGTGTAGAATTCTCTTACTA 3780
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721 TACAGACAGTATCTGTACAGAGTAAGGAGGAGAGGGGGTGTAGAATTCTCTTACTA 3780

Qy  3781 TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCTAGACATCATGGACTT 3840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781 TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCTAGACATCATGGACTT 3840

Qy  3841 TGTCTCTGGGGAGCCAGCACTGTGCAGCACTGTGTGGAACTTCAGTGTCTGAGAGAGTAGGAGAGGCTCCCCTCA 3900
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841 TGTCTCTGGGGAGCCAGCACTGTGCAGCACTGTGTGGAACTTCAGTGTCTGAGAGAGTAGGAGAGGCTCCCCTCA 3900

Qy  3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAAGGGAGAGCCTGGGATGGAGC 3960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAAGGGAGAGCCTGGGATGGAGC 3960

Qy  3961 TTTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAAGAGAGGTGACACCCTTACCCAG 4020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961 TTTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAAGAGAGGTGACACCCTTACCCAG 4020

Qy  4021 TTGTTCAACTCACCCTTCCCTCCCTCGTACCCTCGAGCCCTGATGTGG 4080
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021 TTGTTCAACTCACCCTTCCCTCCCTCGTACCCTCGAGCCCTGATGTGG 4080

Qy  4081 GTGTGAGACGCTCCTCGTTCTCTCTATCCTGCCCATCGGCCCTTCGGGGAGGGAGGAATGT 4140
            |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081 GTGTGAGACGCTCCTCGTTCTCTCTATCGGCCCATCGCCCTTTGGGGAGGGAGGAATGT 4140

Qy  4141 GCCCAAGGACTAAAAAAAGGCCATGGAGCCCAGAGGGGCAACACAGACCTTTCATGG 4200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141 GCCCAAGGACTAAAAAAAGGCCATGGAGCCCAGAGGGGCAACACAGACCTTTCATGG 4200

Qy  4201 GCAAACCTTGGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAATCGATACCCTTCTTC 4260
            ||||||||||||||||||||||
Db  4201 GCAAACCTTGGGGCCCTGCTGTCCCTCGTCACCTGTCACCTTCCAGAGCCAAGGGATCAAAGGAGGA 4260

Qy  4261 TTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGATAGAGAAAGTGA 4320
            |||||||||||
Db  4261 GGAGCCAGGACAGGAGGGAAGTGGGAGGAG------------------------------ 4291

Qy  4321 AAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTGAGTTTACCACTC 4380
            -------------------
Db  4292 ---------------------------------------------------------- 4291

Qy  4381 CCTATCAGTGATAGAGAAAGTGAAAGTGAGTTTACCACTCCCTATCAGTGATAGAGAA 4440
            ----
Db  4292 ---------------------------------------------------------- 4291

Qy  4441 AAGTGAAAGTGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTGAGTTTA 4500
            -----
Db  4292 ---------------------------------------------------------- 4291

Qy  4501 CCACTCCCTATCAGTGATAGAGAAAGTCGAGTTTACCACTCCCTATCAGTGAT 4560
            --------------------------
Db  4292 ---------------------------------------------------------- 4291

Qy  4561 AGAGAAAGTGAAAGTGAGACTCCGTACCAGCAGGACTCCAAATTTAGGCAGCAGGCA 4620
            ||||||||||||||||||||||
Db  4292 ---------------------GGTCCCAGCAGCAGGACTCCAAATTTAGGCAGCAGGCA 4328

Qy  4621 TATGGGATGGATATAAAGGGGCTGTCAGAGATTTCTCAACCCA 4680
            ||||||||||||||||||||||||||||||||||||||||||
Db  4329 TATGGGATGGATATAAAGGGGCTGTCAGAGATTTCTCAACCCA 4388
```

```
RESULT 3
AC099577
LOCUS       AC099577             235759 bp    DNA     linear   HTG 22-FEB-2004
DEFINITION  Mus musculus chromosome 14 clone RP23-171A13 map 14, *** SEQUENCING
            IN PROGRESS ***, 6 unordered pieces.
ACCESSION   AC099577
VERSION     AC099577.3  GI:42734575
KEYWORDS    HTG; HTGS_PHASE1; HTGS FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 235759)
AUTHORS     Birren,B., Nusbaum,C. and Lander,E.
TITLE       Mus musculus chromosome 14, clone RP23-171A13
JOURNAL     Unpublished
REFERENCE   2  (bases 1 to 235759)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Boguslavkiy,L., Boukhgalter,B.,
            Brown,A., Camarata,J., Campopiano,A., Chazaro,B.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
            Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
            Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
            Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
            Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
            Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Liu,G.,
            MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
            McCarthy,M., McEwan,P., McKernan,K., McPheeters,B., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,P.,
            Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
            Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
            Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
            Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
            Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 235759)
AUTHORS     Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
            Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
            Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
TITLE       Direct Submission
JOURNAL     Submitted (22-FEB-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Feb 22, 2004 this sequence version replaced gi:31581779.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -----------  Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
```

```
Qy    4681  GGTAAGAGGGAGTTTCGGTGGTGGGGGCTCTTCACCCACCAGACCTCTCCCCACCTAGAA  4740
            |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    4389  GGTAAGAGGGAGTTTCGGTGGTGGGGGCTCTTCACCCACCAGACCTCTCCCCACCTAGAA  4448

Qy    4741  GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCTGACAGGGTGGCCTT  4800
            |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db    4449  GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCTGACAGGGTGGCCCT  4508

Qy    4801  CCACCAGCCTGGGAAGTTCTCAGTGCCGCAGGAGGTTCACAAGAAAACACTGGATGCCCCT  4860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4509  CCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTCCACAAGAAACACTGGATGCCCCT  4568

Qy    4861  TCCCTTAGGCTGTCTTCTCCATCTTCCTCCTGGGGATGCTCCTCCCCGTCTTGGTTTATC  4920
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4569  TCCCTTAGGCTGTCTTCTCCATCTTCCTCCTGGGGATGCTCCTCCCCGTCTTGGTTTATC  4628

Qy    4921  TTGGCTCTTCGTCTTCAGCAAGATTGCCCGTGCTGTCCACTCCATCTTTCTCTACTGT  4980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4629  TTGGCTCTTCGTCTTCAGCAAGATTTGCCCGTGCTGTCCACTCCATCTTTCTCTACTGT  4688

Qy    4981  CTCCGTGCCTTGCCTTCCTTGCGTCCTTCCTTCCACCCATTTCCACTTCACTTCACC  5040
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4689  CTCCGTGCCTTGCCTTCCTTGCGTCCTTCCTTCCACCCATTTCCACTTCACTTCACC  4748

Qy    5041  TTTTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  5100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4749  TTTTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  4808

Qy    5101  CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  5160
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4809  CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  4868

Qy    5161  CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATGTAAACAATCTTCCA  5220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4869  CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATGTAAACAATCTTCCA  4928

Qy    5221  GTGAGCCACAGCTTCAGTGCTGCTCTGGGTGCTCTCTTCACCCCCTGGCTTGTC  5280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4929  GTGAGCCACAGCTTCAGTGCTGCTCTGGGTGCTCTCTTCACCCCCTGGCTTGTC  4988

Qy    5281  CTGTGTCCATCCTGTCAGGATCCTAGATTGGTCTCCAGCCGCTCGCTACTCCCCTTCCT  5340
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4989  CTGTGTCCATCCTGTCAGGATCCTAGATTGGTCTCCAGCCGCTCGCTACTCCCTTCCT  5048

Qy    5341  GCCGTGTCCTCCTCTCGTCCGTCGCGCCACTGGTGGCCACTGTGGTGGTCGGTCCAC  5400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5049  GCCGTGTCCTCCTCTCGTCCGTCGCGCCACTGGTGGCCACTGTGGTGGTCGGTCCAC  5108

Qy    5401  ATTCTTCAGGATTCTGAAAAGTTAACCAGGTGAGAATGTTTCCCGTGTAGCAGCAGAGA  5460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5109  ATTCTTCAGGATTCTGAAAAGTTAACCAGGTGAGAATGTTTCCCGTGTAGCAGCAGAGA  5168

Qy    5461  TCACGAGTTCTCCCGGAAGTCAGGCTTCCAGCCGTTCAGCCTGCCCAGCTGCCCGGCCA  5520
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5169  TCACGAGTTCTCCCGGAAGTCAGGCTTCCAGCCGTTCAGCCTGCCCAGCTGCCCGGCCA  5228

Qy    5521  CTCTTAGCAAACCTCAGGCACCTTACCCCACATAGACCTTCTGACAGAGAGCAGGCACT  5580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5229  CTCTTAGCAAACCTCAGGCACCTTACCCCACATAGACCTTCTGACAGAGAGCAGGCACT  5288

Qy    5581  TTACATGAGTCCTGGTGGGAGAGCCCATAGGCTACGGTGTAAAAGAGAGGCAGGGAAGTGGT  5640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5289  TTACATGAGTCCTGTGGGGAGAGCCCATAGGCTACGGTGTAAAAGAGAGGCAGGGAAGTGGT  5348

Qy    5641  GGGTGAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACCACCAGAATGACAGACACAGA  5700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5349  GGGTGAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACCACCAGAATGACAGACACAGA  5408

Qy    5701  TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC  5735
            ||||||||||||||||||||||||||||||||||
Db    5409  TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC  5443
```

```
          Web site: http://www-seq.wi.mit.edu
          Contact: sequence submissions@genome.wi.mit.edu
          ------- Project Information
          Center project name: L16015
          Center clone name: 171_A_13
          -------
        * NOTE: This is a 'working draft' sequence. It currently
        * consists of 6 contigs. The true order of the pieces
        * is not known and their order in this sequence record is
        * arbitrary. Gaps between the contigs are represented as
        * runs of N, but the exact sizes of the gaps are unknown.
        * This record will be updated with the finished sequence
        * as soon as it is available and the accession number will
        * be preserved.
        *        1     21715: contig of 21715 bp in length
        *    21716     21815: gap of 100 bp
        *    21816     53733: contig of 31918 bp in length
        *    53734     53833: gap of 100 bp
        *    53834    113500: contig of 59667 bp in length
        *   113501    113600: gap of 100 bp
        *   113601    196948: contig of 83348 bp in length
        *   196949    197048: gap of 100 bp
        *   197049    205662: contig of 8614 bp in length
        *   205663    205762: gap of 100 bp
        *   205763    235759: contig of 29997 bp in length
FEATURES             Location/Qualifiers
     source          1..235759
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
                     /chromosome="14"
                     /map="14"
                     /clone="RP23-171A13"
                     /clone_lib="RPCI-23 Female Mouse BAC"
ORIGIN

Query Match       83.8%;   Score 4804.8;   DB 2;   Length 235759;
Best Local Similarity  92.0%;   Pred. No. 0;
Matches 5316;  Conservative  0;  Mismatches 107;  Indels 357;  Gaps  9;
```

```
Qy        1 GGATCCTGCAAGGTCACACAAGGTCTCCACCCACCAGGTGCCCTAGTCTCAATTTCAGT 60
            ||||||||||||||||| ||||| ||||||||||||||||||||||||||| ||||||||
Db   211665 GGATCCTGCAAGGTCACACAAGAGTCTCCACCCACCAGGTGCCCTAGTCTCAATTTCAGT 211724

Qy       61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTTGGACTTTGTTTTTAT 120
            |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db   211725 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTTGGACTTTGTTTTTAT 211784

Qy      121 TTCAAAAGGGCCTGAATGAGGAGATCTTGTGTCTACCCAGCTCTAAGGGTGCCCGTGA 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211785 TTCAAAAGGGCCTGAATGAGGAGTAGTCTTGTGTCTACCCAGCTCCAAGGGTGCCCGTGA 211844

Qy      181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTGAGACTCTTCTTCTCGTGAGGTGGAAGCAGAATAAAGCAATTT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211845 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 211904

Qy      241 TCCTTAAAGCCAAAATCCTGGAGCCTCTAGACTCTTCTTCTCGGGCTCT 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211905 TCCTTAAAGCCAAAATCCTGCGCCTCTAGACTCTTCTTCTCTGACCTGGGCTCT 211964

Qy      301 AGGGTGGGAGGTGGGGCTTGGAAGAAGAAGGTGGGAAGTGGCAAAGCCGATCCCTAG 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211965 AGGGTGGGAGGTGGGGCTTGGAAGAAGAAGGTGGGAAGTGGCAAAAGCCGATCCCTAG 212024

Qy      361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCCAGCC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212025 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCCAGCC 212084

Qy      421 AAACATAGCAAGAAGTGATACCTTCCTTTGACTTCCCAGGGCCAGTACCTGTCAGTT 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212085 AAACATAGCAAGAAGTGATACCTTCCTTTGTGACTTCTCCAGATCAGACTCGAGCTGC 212144

Qy      481 GAAACAGGATTTAGAGAGCCTCGAACTCACCTGAACTCTGAAGCTCATCCACCAAGCA 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212145 GAAACAGGATTTAGAGAGCCTCGAACTCACCTGAACTCTGAATGATTAGCAGAGCTGGGCCAC 212204

Qy      541 AGCACCTAGGTGCCACTGCTAGTAGTATCCTACGCTGATAATGCAGAGCTGGGCCAC 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212205 AGCACCTAGGTGCCACTGCTAGTAGTATCCTACGCTGATAATGCAGAGCTGGGCCAC 212264

Qy      601 AGAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGTATGACCCCC 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212265 AGAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGGTATGACCCCC 212324

Qy      661 TCAGCAGATGAGTAATGTCCCCTTAGATCCCATCCCAGATCCCATCCCCTAGATCCCATCCCAGACACG 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212325 TCAGCAGATGAGTAATGTCCCCTTAGATCCCATCCCAGATCCCTAAGAGGACATG 212384

Qy      721 GGATGAGAGATGTAGTCATGTGGCATTCAAAACAGCTATCCACAGTGTCCCTTGCCC 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212385 GGATGAGAGATGTAGTCAAGTGACATTCAAAACAGCTATCCACAGTGTCCCTTGCCCC 212444

Qy      781 TTCCACTTAGCCAGGAGACAGTAACCTTAGCCTATCTTCCTCCCCATCTCCCAG 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212445 TTCCACTTAGCCAGGAGACAGTAACCTTAGCCTATCTTCCTCCCCATCCCCCAG 212504

Qy      841 GACACACCCCCTGGTCTGCAGTATTCATTTCTTCCTTCACGTGCCCCTCTGTGACTTCCAT 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212505 GACACACCCCCTGGTCTGCAGTATTCATTTCTTCCTTCACGTCCCCTCTGTGACTTCCAT 212564

Qy      901 TTGCAAGGCTTTTGACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212565 TTGCAAGGCTTTTGACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 212624

Qy      961 GCCATCTCAAGAGAAAGCAGACAACAGCAGGGGACCAGATTTTGGAAGGATCAGGAACTAAA 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212625 GCCATCTCAAGAGAAAGCAGACAACAGCAGGGGACCAGATTTTGGAAGGATCAGGAACTAAA 212684

Qy     1021 TCACTGGCGGGGCCTGGGGTGGGGTAGAAGAGAGTGAGTGAGTCCGTCCAGCTAAGCCAAGC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212685 TCACTGGCGGGGCCTGGGGTGGGGTAGAAAAAGAGAGTGAGTGAGTCCGTCCAGCTAAGCCAAGC 212744

Qy     1081 TAGTCCCGAGATACTCTGCACAGTCCAGAGCTTTAGGAATGTGGGTC 1140
            ||||||||||||||||||||||||||||||||||||||||| ||||||
Db   212745 TAGTTCCCGAGATACTCTGCACAGTCCAGAGCTTTAGGAATGTAGGTC 212804

Qy     1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCACCTACAGAC 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212805 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCACCTACAGAC 212864

Qy     1201 ACACTCGTGTGTGGCCAGACTCCTGTCGTTGTCAACAGCGCCCTGTGTTCTGACCACTGAGCTAG 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212865 ACACTCATGTGTGGCCAGACTCCTGTCGTTGTCAACAGCGCCCTGTGTTCTGACCACTGAGCTAG 212924

Qy     1261 GCAACCAGAGCATGGGCCCTGTGTCGTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212925 GCAACCAGAGCATGGGCCCTGTGTCGTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG 212984

Qy     1321 CAGGGGAGGGAGAACAGAGAACAGAGAACAGAATAAGGAAGGAAGAGGAAGGCCAGTCAATCAGA 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212985 CAGGGGAGGGAGAACAGAGAACAGAGAACAGAATAAGGAAGGAAGAGGAAGGCCAGTCAATCAGA 213044

Qy     1381 TGCAGTCAGAAGAATGGGAAGTCCAACACACCAACCAACCACAGAGTTGAGCAGCAGAAACAGAAAGGGAG 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   213045 TGCAGTCAGAAGAATGGGAAGTCCAACACCAACCACCACAGAGTTGAGCAGCAGAAACAGAAAGGGAG 213104

Qy     1441 AGATTCTGGCATAAGGAGAGCCCAGGCCCCCCAAGTCTCCCTTT 1500
            ||||||||||||||||||||||||||||||||||||||||||||
Db   213105 AGATTCTGGCATAAGGAGAGCCCAAAGAGAGCCCAGGCCCCCCAAGTCTCTCCTTTT 213164

Qy     1501 ATACCCTCATCCCGTCCCAATTAAGCCACTCTTCCTAGATCAGACTCGAGCTGC 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   213165 ATACCCTCATCCCGTCTTCCAATTAAGCCCACTCTTCCTAGATCAGACTCGAGCTGC 213224

Qy     1561 AGCGAAGAGACCCGTAGGAGGATCACACTGAGGGAGATGTGTGACTCCCAGG 1620
```

```
Qy    2697 GCCCCACCCACCATAAGGGAGTGAACTATCCTAGGGGCTGGCGACCTTGGGGACACA 2756
           ||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db  214352 GCCCCACCCACCATAAGGGAGTGAACTATCCTAGGGGCTGGCGACCTTGGGGACACA 214411

Qy    2757 CCACATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCCGTCGTCCTGCCCACCTCC 2816
           ||||| |||| |||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  214412 CCATATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCCGTCGTCCTGCCCACCTCC 214471

Qy    2817 ACACTCTAGAGCTATATTGAGAGGTAGACAGTAGAGGGTGGGAGTGGTAGCAGGGAGA 2876
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214472 ACACTCTAGAGCTATATTGAGAGGTAGACAGTAGAGGGTGGGAGTGGTAGCAGGGAGA 214531

Qy    2877 GTGTTCCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGTCTCC 2936
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214532 GTGTTCCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGCCTCC 214591

Qy    2937 TGAACACAATGTCTACTTAGTTATAAACAGGCATGAACCTGCTAAAGACCCAACATCTACGA 2996
           |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  214592 TGAACACAATGTCTACTTAGTTATAAACAGGCATGAACCTGCTAAAGACCCAACATCTACGA 214651

Qy    2997 CCTCTGAAAAGACACAGCCCTGGGAGGACAGCAGGGGTTGTCTCTGAGCCTTGGGTGCTTGAT 3056
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214652 CCTCTGAAAAGACACAGCCCTGGGAGGACAGCAGGGGTTGTCTCTGAGCCTTGGGTGCTTGAT 214711

Qy    3057 GGTGCCACAAAGGAGGGCATGAGTGTGAGTATAAGGCCCCAGGAGCGGTTAGGAGAAGGGCA 3116
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214712 GGTGCCACAAAGGAGGGCATGAGTGTGAGTATAAGGCCCCAGGAGCGGTTAGGAGAAGGGCA 214771

Qy    3117 CTTGGGAAGGGGTCAGTCTGCAGAGCCCCTATCCATGGAATCTGAGAGTGAGGGTTGCAGGGGCCAACT 3176
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214772 CTTGGGAAGGGGTCAGTCTGCAGAGCCCCTATCCATGGAATCTGAGAGTGAGGGTTGCAGGGGCCAACT 214831

Qy    3177 GGTGTAAATCTCTGGCCTGCCAGGCATTCAAAGCACCTGCATCCTCTGGCAGCCTG 3236
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214832 GGTGTAAATCTCTGGCCTGCCAGGCATTCAAAGCACCTGCATCCTCTGGCAGCCTG 214891

Qy    3237 GGGAGGCGGAAGGGAGGAGCAACCCCCACTTATACCCTTCTCCCTAGCGCCCCAGGATTAAC 3296
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db  214892 GGGAGGCGGAAGGGAGGAGCAACCCCCACTTATACCCTTCTCCCTAGCGCCCCAGGATTAAC 214951

Qy    3297 ACCTCTGGCCTTCCCCCTTCCCCATCAGAGTGGGTTGCAGAGTGAGGGTTGCATGTCGTGTGTG 3356
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214952 ACCTCTGGCCTTCCCCCTTCCCCATCAGAGTGGGTTGCAGAGTGAGGGTTGCATGTCGTGTGTG 215011

Qy    3357 AAAACCTACATGTCCAAAACATCATGTGCACGATATGATCAGTATGTGCAAGATATGTGCAGGGTA 3416
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  215012 AAAACCTACATGTCCAAAACATCATGTGCACGATATGATCAGTATGTGCAAGATATGTGCAGGGTA 215071

Qy    3417 GAAAGGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTGTG 3476
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  215072 GAAAGGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTGTG 215131

Qy    3477 TCTGACTGAAAACGGGCATGGGCATGGGCATGGGGTTCAGTTCTGAGGTTACCAGAC 3536
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  215132 TCTGACTGAAAACGGGCATGGGCATGGGCATGGGGTTCAGTTCTGAGGTTACCAGAC 215191

Qy    3537 GTCTCAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCGTCTAGAGATGTGGCAACAG 3596
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  215192 GTCTCAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCGTCTAGAGATGTGGCAACAG 215251

Qy    3597 AGATTGGATGATGGCCTGCTATCTCAAGGGACCATGGAAAAATAGAATGGACACTCTATATGT 3656
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  215252 AGATTGGATGATGGCCTGCTATCTCAAGGGACCATGGAAAAATAGAATGGACACTCTATATGT 215311

Qy    3657 GTCTCAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCGTCTAGAGATGTGGCAACAG 3716
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  215312 GTCCTCAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCGTCTAGAGATGTGGCAACAG 215371

Qy    3717 AGACTACAGACAGTATCTGTACAGATGTAAGGAGAGAGAGGGG-GTGTAGAATTCTCT 3775
           ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  215372 AGACTACAGACAGTATCTGTACAGATGTAAGGAGAGAGAGAGGAGGGGGCGTAGAATTCTCT 215431
```

```
Db  213225 AGGCGAAGAGACCCGTAGGGAGGATCACACTGGATGAAGGAGATGTGTGGAGAAGTCCAGG 213284
Qy    1680                                                               
Db  213285 G-AACCTAAGAGCCCAGAGCCAGAGCAAGACCTAAAGAGCAAGATAAAGGTGTGGCGCAGG 213343

Qy    1681 CTGTGCACACAGAGGTCGAGGACTGGTGTGGAGAGCCTCAAGATAAGGATGATGCTCAGA 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213344 CTGTGCACACAGAGGTCGAGGACTGGTGTGGAGAGCCTCAAGATAAGGATGATGCTCAGA 213403

Qy    1741 ATGGCGGGGGGG-GGGGATTCTGGGGGAGAGGAGAGAGGTGAAGAGGAGCCTGAGAACA 1799
           ||||||||||   |||||| |||||||||||||||||||||||||||||||||||||||
Db  213404 ATGGCGGGGAGGGTGGGGGAGGGTCTGGGGGAGAGGAGAGAGGTGAAGAGGAGCCTGAACA 213463

Qy    1800 GAGAATCTGGAAGCGCTGAAACCATACCATAAAGGGAAGAACCCAGGCTACCTTTAGAT 1859
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213464 GAGAATCTGGAAGCGCTGAAACCATACCATAAAGGGAAGAACCCAGGCTACCTTTAGAT 213523

Qy    1860 GTAAATCATGAAAGAAGGAGAAGGAAGCTGGAGAGAGT---AGAAGGACCCCGGGGGC 1916
           ||||||||||||||||||||||||||||||||||||||   |||||||||||||||||
Db  213524 GTAAATCATGAAAGAAGGAGAAGGAAGCTGGAGAGAGTAGAGAAGGACCCCGGGGGC 213583

Qy    1917 AAGACATGAAGCAAGGACAAGCAGCCCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGAAGGC 1976
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db  213584 AAGACATGAAGCAAGGACAAGCAGCCCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGAAGGC 213643

Qy    1977 AGAGCCCTGTATGAGACCAGAACAGCAGACCAGAGACCAGAGACAGGAACA 2036
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db  213644 A-----------AGCACCAGAACAGCAGACCAGAGACCAGAGACAGGAACA 213691

Qy    2037 GAAGGTAGAACAGGAACAGAGACACAGGACGGGGGAGCTAGGGTAACAAAGGAATGGTCCTTC 2096
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213692 GAAGGTAGAACAGGAACAGAGACACAGGACGGGGGAGCTAGGGTAACAAAGGAATGGTCCTTC 213751

Qy    2097 TCACCTGTGGCCAGAGCGTCCAGCGTCATCTGTGTCCACATACTCTGACATGTTCATCAGACTGCA 2156
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213752 TCACCTGTGGCCAGAGCGTCCAGCGTCATCTGTGTCCACATACTCTGACATGTTCATCAGACTGTA 213811

Qy    2157 GGGCTGGCTTGGGAGGCAGCTGGAAAGATGTGAGAGCCAGGGCGGAGCAAGGGGGCCT 2216
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213812 GGGCTGGCTTGGGAGGCAGCTGGAAAGATGTGAGAGCCAGGGCGGAGCAAGGGGGCCT 213871

Qy    2217 AGGAAAGGAAGAAGAGGAAGGGCAAACCAGGGCAAACCAGGGCAGAGCCCAGAACTGAGTTA 2276
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213872 AGGAAAGGAAGAAGAGGAAGGGCAAACCAGGGCAAACCAGGGCAGAGCCCAGAACTGAGTTA 213931

Qy    2277 ACTCCTTCCTTCGTTGTGCATCTGTCTGACCCAGGAGGCAGTGGAACTCTGTGACCACCATCCCCC 2336
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213932 ACTCCTTCCTTCGTTGTGCATCTTGCATCTGTCTGACCCAGGAGGCAGTGGAACTCTGTGACCACCATCCCCC 213991

Qy    2337 ATGAGCCCCCACTACCCATACCCATACCCAGCCAGCAGTGACATATATTAGGGCCATGGGCTAACCCCTGG 2396
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  213992 ATGAGCCCCCACTACCCATACCCATACCCAGCCAGCAGTGACATATATTAGGGCCATGGGCTAACCCCTG 214051

Qy    2397 ATCCCCAAAGAGCAACCATTGGCATAGGTGGCTGCAAATGGGAATGCAAGGTTGAATCA 2456
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214052 ATCCCCAAAGAGCAACCATTGGCATAGGTGGCTGCAAACGGGAATGCAAGGTTGAATCA 214111

Qy    2457 GAATGCCCTTCAAGAATACTGCAATGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTTAAAG 2516
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214112 TCAATGCCCTTCAAGAGTCTTGGTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTTAAAG 214171

Qy    2517 AGCTTGCACACAGAGCCTCAAGTGACCTGCAGAGCAGTGACCTGCAGAGCAGTGGCAACAC 2576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214172 AGCTTGCACACAGAGCCTCAAGTGACCTGCAGAGCAGTGACCTGCAGAGCAGTGGCAACAC 214231

Qy    2577 GTCTCAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCGTCTAGAGATGGGCAACAG 2636
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  214232 GTCCCTCAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCGTCTAGAGATGGGCAACAG 214291

Qy    2637 GTCCCCTTCAAGAATACTGCCATGCCAAGTAAGATACTGCTTAGGAGGGTATGCTCCT 2696
           ||||||||||||||||||||| |||| |||||||||||||||||||||||||||||||
Db  214292 GGTCCCCTTGGAGAATACTGCCATGCCAGTAGCTAAGACCTTAGGAGGGTATGCTCCT 214351
```

Qy   3776 TACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCTAGACAATCAT 3835
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215432 TACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAATTGGATGCCTCCCTAGACAATCAT 215491

Qy   3836 GACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGTCTGAGAGTAGGAGGCTCC 3895
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215492 GACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGTCTGAGAGTAGGAGGCTCC 215551

Qy   3896 CCTCCAGCCTGAAGCTATGCAGATAGCCAAGGGGTTGAAAGGGGAGGGGAGAGCCTGGGATG 3955
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215552 CCTCAGCCTGAAGCTATGCAGATAGCCAAGGGGTTGAAAGGGGAGGGGAGAGCCTGGGATG 215611

Qy   3956 GGAGCTTGTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAAGAGAAGGTCACCCTTA 4015
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215612 GGAGCTTGTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAAGAGAAGGTGACCCTTA 215671

Qy   4016 CCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGCCTCGGGTAGGGG 4075
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215672 CCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGCCTCGGGTAGGGG 215731

Qy   4076 AGGTGGTGTGAGACGCTCCTGTCTCCTGCATGCCGCCTGCCTTTGGGGAGGAGG 4135
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db 215732 AGGTGGTGTGAGACGCTCCTGTCTCCTCCATCGGCCCTTTGGGGAGGAGG 215791

Qy   4136 AATGTGCCCAAGGACTAAAAAAAGGCCATGGAGCCCAGAGGGGCCAGGGCAACAGACCTTT 4195
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215792 AATGTGCCCAAGGACTAAAAAAAGGCCATGGAGCCCAGAGGGGCCAGGGCAACAGACCTTT 215851

Qy   4196 CATGGGCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAATCGATACCCT 4255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 215852 CATGGGCAAACCTTGGGCCCGTCCTGCTCACTCCTGTCCCTGTCCTATCAGTGTGATAGAGAAAG 215911

Qy   4256 TCTTCTTCTAAACGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGTGGGAGGA 4315
           ||||||||||||||||||||
Db 215912 GAGGAGGAGCCAGACAGGAGGGAAGTCGGGAGGA 215946

Qy   4316 AGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTCGAGTTTAC 4375
           ||||||||||||||
Db 215946

Qy   4376 CACTCCCTATCAGTGATAGAGAAAGTCGAGTTTACCACTCCCTATCAGTGATA 4435
           ||||
Db 215946

Qy   4436 GAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTAGTAGAGAAAGTGAAAGTCGA 4495
           ||||||||||||||||||
Db 215946

Qy   4496 GTTTACCACTCCCTATCAGTGATAGAGAAAGTGAGTTTACCACTCCCTATCA 4555
           ||||||||||||||
Db 215946

Qy   4556 GTGATAGAGAAAGTGAAAGTCGAGTCGAGTTTACCACTCCCTAAATTTAGGCAGC 4615
           |||||||||||||||||||
Db 215947 GGGTCCCAGCAGCAGGACTCCCAAATTTAGGCAGC 215979

Qy   4616 AGGCCATATGGGATGGGATATAAAGGGCGGCTGGGATATAAAG---CAGAGATTTC 4671
           |||||||||||||||||||||||||||||||||||||||||| ||||||||
Db 215980 AGGCCATATGGGATGGGATATAAAGGGCTGGGACTGCAGAGCTGTCAGAGAGATTTC 216039

Qy   4672 TCCAACCCAGGTAAGGGAAGTTCGGGTGGGAAGTTCGGGTGGGAAGACCCTCCCC 4731
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216040 TCCAACCCAGGTAAGAGGGGAAGTTCGGGTGGGAAGTTCGGGTGGGAGCACCTCTCCCC 216099

Qy   4732 CACCTAGAAGGGAAGCTGCCTTTCCTGGAAGTGGGGTTCAGGCGGGTTCAGGGCCGGTCAGAGATCTGACAG 4791
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216100 CACCTAGAAGGAGAACTGCCTTTCCTGGAAGTGGGGTTCAGGCGGGTTCAGGCCCGGTCAGAGATCTGACAG 216159

Qy   4792 GGTGGCCTTCCACCAGCCTCAGTGCCAGGAGGTTTCCAAGAAACACTG 4851
           ||||||||||||||||||||||||||||||||||||||||||||
Db 216160 GGTGGCCTTCCACCAGCCTCAGTGCCAGGAGGTTTCTCAGTGGC---AGGTTTCCACAGAAACACTG 216216

Qy   4852 GATGCCCCTTCACGCTGTCTTCTCCATCGTCTTCCTCATCTTCCTCCCTGGGGGATGCCTGTGTCTGTCCACTCCATCTTT 4911
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216217 GATGCCCCTTCACGCTGTCTTCGTCTTCCTCCATCTTCCTCAGCAAGAGATTGCCCGTGTCTGTCCACTCCATCTTCT 216276

Qy   4912 TGGTTTATCTTGGCTCTTCGTCGTCCTTGCCTTCCTTCAGCAAGATTGCCCGTGTCTGTCCACTCCATCTTT 4971
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216277 TGGTTTATCTTGGCTCTTCGTCGTCCTTGCCTTCCTTCAGCAAGATTGCCCGTGTCTGTCCACTCCATCTTT 216336

Qy   4972 CCTTACTGTCTCCGTGCCTTGCCTTCCTTCGTGTGCTTGCGTGTCCTTCCTTCACCCATTTCT 5031
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216337 CCTTACTGTCTCCGTGCCTTGCCTTCGTCCTTGCGTGTCCTTCCTTCACCCATTTCT 216396

Qy   5032 CACTTCACCTTCTCCCCTTCCTCATTTGTATTCATTCATCCTTCCTTCCTT------- 5083
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db 216397 CACTTCACCTTTCTCCCCTTCCTCATTTGTATTCATCCTCCTTCCTTCCTTCCTTCCTT 216456

Qy   5084 ----------------------------CCTTCCTTCCTTCCTTCCTTCCTT 5111
                                       |||||||||||||||||||||||
Db 216457 CCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTT 216516

Qy   5112 CCTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCTGTGTCAGAG 5171
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db 216517 CCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTCTGTGTCAGAG 216576

Qy   5172 TGCTGAGAATCACACCTGGGGTTCCCACCCTTATGTAAACAATCTTCCAGTGAGCCACAG 5231
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216577 TGCTGAGAATCACACCTGGGGTTCCCACCCTTATGTAAACAATCTTCCAGTGAGCCACAG 216636

Qy   5232 CTTCAGTGCTGCTGGGTGCTCTCTTCTCTTCTTACCTTCTCCACCCCGTGTCCTCCTGCCTGTCCTC 5291
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216637 CTTCAGTGCTGCTGGGTGCTCTCGATTGGTCTCCACCCCGTGTCCTCCTGCCTGTCCTC 216696

Qy   5292 TGGTCAGATCTCTAGATTGGTCTCCACCGTGTGCTGAGAATGTTCCCCTGTAGACAGCAGATCACGATTCTC 5351
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216697 TGGTCAGGATCTCTCGGATTGGTCTCCACCGTGTGCTGAGAATGTTCCCCTGTAGACAGCAGATCACGATTCTC 216756

Qy   5352 TCTCTGTCCAGTCGCGCACTGTGGTGCCACTGTGGTGTCGTTCCCAGCTGTGGTGTCGTTCCTC 5411
           ||||||||||||||||||||||||||||||||||||||||||||||
Db 216757 TCTCTGTCCAGTCGCGCACTGTGGTGCCACTGTGGTGTCCACATTCTTCAGGA 216816

Qy   5412 TTCTCTGAAAGTTAACCAGTGAGAATGTTCCCCTGTAGACAGCAGATCACGATTCTC 5471
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db 216817 TTCTCTGAAAGTTAACCAGTGAGAATGTTCCCCTGTAGACAGCAGATCACGATTCTC 216876

Qy   5472 CCGGAAGTCAGGCTTCCAGCCCTCTTCTTCTGCCCAGCTGCCCGGCCACTCTTAGCAAA 5531
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216877 CCGGAAGTCAGGCTTCCAGCCCTCTTCTTCTGCCCAGCTGCCCGGCCACTCTTAGCAAA 216936

Qy   5532 CCTCCAGGCACCCCTTACCCCACCATAGACCCTGACAGAGAGAGCAGCAGGAGTCACTGGAGT 5591
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216937 CCTCAGGCACCCCTTACCCCACCATAGACCCTGACAGAGAGAGCAGCAGGAGTCACTTTACATAGAGT 216996

Qy   5592 CCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGAGGCAGGAAGTGGTGGTGTAGGAAA 5651
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 216997 CCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGAGGCAGGAAGTGGTGGTGTAGGAAA 217056

Qy   5652 GTCAGGACTTCACATAGAGCCTAGCCACCCAGAAATGACAGACAGATCCCTCCTATC 5711
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 217057 GTCAGGACTTCACATAGAGCCTAGCCACCCAGAAATGACAGACAGATCCCTCCTATC 217116

Qy   5712 TCCCCCATAAGAGTTTGAGT 5731
           ||||||||||||||||||||
Db 217117 TCCCCCATAAGAGTTTGAGT 217136

REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
  1 (bases 1 to 185702)
AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,G., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,A., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,W., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,J., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,M., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,J.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olanrpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,M., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Taylor,C.,
            Steimle,M., Strong,R., Sutton,A., Svatek,R., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,J., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,M., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 185702)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (15-AUG-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 185702)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 15, 2002 this sequence version replaced gi:23195973.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,

individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contig within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
------------- Genome Center
    Center: Baylor College of Medicine
    Center code: BCM
    Web site: http://www.hgsc.bcm.tmc.edu/
    Contact: hgsc-help@bcm.tmc.edu
------------- Project Information
    Center project name: KCAL
    Center clone name: CH230-249H16
------------- Summary Statistics
    Assembly program: Phrap; version 0.990329
    Consensus quality: 158880 bases at least Q40
    Consensus quality: 159964 bases at least Q30
    Consensus quality: 160601 bases at least Q20
    Estimated insert size: 162428; sum-of-contigs estimation
    Quality coverage: 9x in Q20 bases; sum-of-contigs estimation
----
* NOTE: Estimated insert size may differ from sequence length
*   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
*   consists of 1 contigs. Gaps between the contigs
*   are represented as runs of N. The order of the pieces
*   is believed to be correct as given, however the sizes
*   of the gaps between them are based on estimates that have
*   provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*   1 185702: contig of 185702 bp in length.

FEATURES   Location/Qualifiers
     source   1..185702
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-249H16"
     misc_feature   2778..83571
                     /note="clone_boundary
                     clone_end:T7
                     site:
                     end sequence:BZ214368"
     misc_feature   180811..182798
                     /note="wgs_contig"
     misc_feature   184218..185702
                     /note="wgs_contig"

ORIGIN

Query Match   40.9%;   Score 2343;   DB 2;   Length 185702;
Best Local Similarity   71.2%;   Pred. No. 0;
Matches 4321;   Conservative   0; Mismatches 795;   Indels 951;   Gaps 55;

Qy   1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCACCAGGGTGCCCTAGTCTCAATTCAGT 60
            |||||| |||||||||||||||||||||||||||||
Db   97021 GGATCCTCAAGGTCACACAAGGGTCTCCACCACCAGGGTGCCCTAGTCTCAACTTCAGT 97080

Qy   61 TTCCATGCCTTGTTCTCACAATGCTGGGCCTCCCCAGAGAGCTAATTTGGACTTTGTTTTAT 120
            |||||| ||||||||||||||||||||||||||||||
Db   97081 TTCCATGTGCTTCCCTACAATGTTGGCCTCCCCCAGAGAGCTAATTTTGAC-TTGGTTTTAT 97139

Qy   121 TTCAAAAGGCCTGAATGAGAGTAGATCTTGTGTACCCAGCTCTAAGGGGTGCCCGTGA 180
            ||||||| ||||||||||||||||||||||||||||
Db   97140 TTCAAAAGGCCTGAATGAGAGTAGATCTTGTGTACCCAACCCTAAGGATGCCCGTGA 97199

Qy   181 AGCCCTCAGACCTGGACCTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 240
            |||||| ||| ||||| |||||||||||||||||
Db   97200 AGCCCTGAGACCTGGAGCCCTTTGAAACAGCACAGCACCTTAGGCAGCAGAAACACAATAAAGCAATTT 97259

Qy   241 TCCTTAAAGCCAAAATCCTGCCTCTCAGAGACTCTTCTTCTGACCTCGGTCCCT-GGGCTC 299

Db 97260 TCCTTCAAGCCAAATCCTGTCTCTTCCTTCACTGACCTCGGTCCGTCGGGGCTC 97319
       |||||| ||||||||||||| ||||||||||||
Qy   300 TAGGGTGGGGAGTGGGGCTTGGGAGAAGGTGGGAAGTGGCAAAAGCCGATCCCTA 359
       |||||||||||||||||| |||| ||||||||||||||||
Db 97320 TAGGGTGGGGAGTGGGGACTTG---AAAGAAGGAGGGAAGTGGCAAAGGCCACTCCCCA 97376
       ||||||||||||||||||||
Qy   360 GGGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCTCCAGC 419
       |||||||||||||||||||||||||||||||||||||
Db 97377 GGACTCTGTGAAGTTCTGAGCCTTCCCAGTACAGTGCTGGCTCATAGATCCTCCTCCAGC 97436
       |||||||||||||||
Qy   420 CAAACATAGCAAGAAGTGATACCTCTTGTGACTTCTTGGTACCTGTCAGGT 479
       ||||| ||||||||||||||||||||||
Db 97437 CAAACACAGCAAGAAGTGATGCTTCCTTGTGACTTCCCCAGTACCTGTCCGGT 97496
       |||||||||||||||||
Qy   480 TGAAACAGGATTTAGAGAGCCTCGAACTCACCTGAAGCTCATCACCAAGC 539
       ||||||||||||||||||||||||||||||
Db 97497 TGAAACAGGATTTAGAGAAGCCTCTAAACTCACCTGAGCCTCGAGCCCATCTAC----C 97552
       ||||||||||
Qy   540 AAGCACCTAGGTGCCACTGCTAGTAGTATCCTACGCTGATATATGCAGAGCTG--GGC 597
       |||||| |||||||
Db 97553 ACACACCTACATGCCGGTTCTAATCAGTATCCTACACTGATAATACGCAGAGCTGACCAC 97612
       |||||
Qy   598 CACAGAAGTCCTGGGGGTGTAGGGGTGAGGAACTGACCAGTGACTTTTCAGTCGG 657
       |||||||| |||||||||||
Db 97613 GGGAGAAGGGCTGGAGTGTAGAGAACTGGCCAGTGACTTTTCAGTCTGCAGAGGTATGACC 97672
       |||
Qy   658 CCCTCAGCAGAT---GTAGTAATGTCCCCTTAGATCCCATCCCAGGCAGGTCTCTAAGAG 714
       |||| |||||||
Db 97673 CCCTTGGCAGATATGGCAGCAGATGGGTCCCCTTAGAACCCATCCCAGGCAGGTCTCTGAGAG 97732
       |||
Qy   715 GACATGGGATGGAGAGATGTAGTCATGTGGCATTCCAAACACAGCTATCCCACAGTGTCCCT 774
       ||||||||
Db 97733 GACACAGGATGGACAGATGGTAGTCATGTGGCATTCCAAACACAGCTATCCCACAGTGACTCT 97792
       ||||
Qy   775 TGCCCCTTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTCTTCTTCCCATCC 834
       ||||||||||||||||
Db 97793 TGTCCCCTTCCACTTAGCCATGGAGGATAGCAACCTTAGCCTACCTTTCTTCTTCTTCCCATCT 97852
       |||
Qy   835 TCCCAGGACACACCCCCTGGTCTGCAGTATTCATTTCCTTCACGTTCCCCTCTGTGAC 894
       |||||||||||||||
Db 97853 TCCCAGGACCCAACCCCGTCTATGGTATTCATTTCATTTCTT-CTTCATTTCCCCCACCATGAC 97911
       |||
Qy   895 TCCCATTTGCAAGGCTTTTGACCTTCTGCAGCTGGCAGTAGTAGAGAGATGAGTTTGGCTCTTATATG 954
       |||||
Db 97912 TTTCCATTTGCAGGGCTTTTTTACCCTCGCAGTAGTGGAAGATGGAGTTTGGCTCTTATATG 97971
       |||||||||
Qy   955 TGGCTCTGAAAGACAATGGGGATTGGAGACATCTTCTTTGAGTCTCCCCCCACCT 1014
       ||||||||||||||||
Db 97972 TGGCTCTGAAAGACAATGGGGATTTGAAAACATCTTCTTTGAACCTGTTTGTTAGG-GGGGAC 98030
       |||||||||||||||
Qy   1015 ACTAAATCACTGGGGGGCCTGGGGGTAGAAAAAGAGTGAGTCGCTCCAGCTAAG 1074
       ||||||||||
Db 98031 ACTAAATCACAGGAGGGTCTGGAGTTCATGAAAGTGAAAAGTCA-------GCTCCAGCTAAG 98082
       |||||||||
Qy   1075 CCAAGCTAGTCCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATG 1134
       |||||||||
Db 98083 CAAAGCCAGATCCTTGGGTATT----CACAGCTGGGCTGCT-GAGGTAGGTTTAGGAATG 98137
       |||||
Qy   1135 TGGGTCTGAAAGACAATGGGATTGGAGACATCTTCTTTGAGTCTCCCCCTCAACCCCACCT 1194
       ||||||||||||||
Db 98138 TGGCTCTGAAAGACAATGGGATTTGAAAACATCTTTGAACCTTGCCCTGCTCTGGGCT 98197
       |||
Qy   1195 ACAGACACACTGTGTGGCCCAGCCCTCTGTGTTCTGACCACTG 1254
       ||||||||||
Db 98198 AAATCTCCCCTCAACCCCACCCATAGAC--------ACGCTTGTGCATGGCCCAGACTG 98247
       |||
Qy   1255 AGCTAGGCAACCAGCATGGGCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAA 1314
       ||||||||||
Db 98248 AGTTAGGCAACTGGCAGCGGCCCCTGTG-TGAGGACGAGG----GGTTGCCAATAGCAA 98302
       |||||
Qy   1315 AAACAGCAGGGAGGAGGAACAGACAGAGAAATAAGGAAGAAGGAAAAGGCCAGTCA 1374

       ||||| ||||||| || ||||||||||||||||||

Db 98303 AAACAGCAGGTGAGGAGAACAGAAAATGAAAATAAGGAAGAAGACGGAACGGCC---CA 98359
       |||||||||||||||||||| ||||||||||
Qy   1375 ATCAGATCGAGTCAGTCAGAAGATGGAAGCCAACACACAGCTTGAGCAGGAGGAAACAGAAA 1434
       |||||||||||||||||||||||
Db 98360 ATCAGATGCAGTCAGTCAGAAGATGGAAGCCAACCAATGCATAGCTTGAACAAGAAAATAGAAG 98419
       |
Qy   1435 AGGGAGAGA--------------- 1443
       ||| ||||
Db 98420 AGGAAGAGATGGCCAGATCATTAGGGTGGACTTCCTCCATTAATTAAGATGCCTCTGTGT 98479
       ||||||||||||||||||||||||||||
Qy   1444 -------TTCTGGCATAAGGCCACAGAAAGAAGAGCCCAGGCCC 1484
       |||||||||||||||||||
Db 98480 CCGACTCTACATCCCTGGGTACTGTCATTGCAACTCAGGAATAGCAGACACAGAGAACT 98539
       |||||
Qy   1485 CCCAAGTCTCCTCTTTATACCCTCATCC- 1513
       |||||||||||||||||||||
Db 98540 GTATATGGCCATTTTCACATTTCAATCCTGAAGCTTGACACACTTGACAATTTCTAGTT 98599
       ||||||
Qy   1514 ----------------------------- 1513
       ||
Db 98600 TACATTGTGTTACTCTATACAGATATGTTCTTCATGGACAGTCGTGTGATGCTCTCCCTGAAT 98659
       ||||||
Qy   1514 ----------------------------- 1513
       |||
Db 98660 ACAAGGTCCTTTTTTTATGAAATCAGTGTGCATTAAATCAGTCTGCATAGGGAGGGCCA 98719
       |||
Qy   1514 ----------------------------- 1551
       |||||
Db 98720 CAGGAAGAGCCCAGGTCCCTCTTTATGCCCTTCCCTCCCTAAGCCCATTC 98779
       ||||||||||||||||
Qy   1552 CTGAGCTGCAGCGAAGAGACCCGTAGGGAGGGATCACACTGATGAAGGAGATGTGGAG 1611
       |||||||||||||
Db 98780 CTCTTC-CTGGAACAGATCTGAGCGACTGGACCGAGAGAAATCCTAGAGTCTAAAGAGGATCA 98838
       |||||||||||||||||
Qy   1612 AAGTCCAGGCAACCTAAGACCAAGAGCCAGAGGATAAAGAGGTCGCTTCAAAG 1671
       |||||||||
Db 98839 CAGTGGAAGAGATGTGTGACCAGAGCCTAAAGAGCAAGAGAGTAAAGAGCCAGCAGCAAAG 98898
       |||||
Qy   1672 GTGGCCAGGCTGTG--CACACAGAGGGTCGAGGACTGGTTGGTAGAGCCTCAAGATAAGGA 1729
       |||||||||||
Db 98899 GCAGCTAGCCCGTGCACACACAGAGGGTCAAGGACTAGCCGGTAGAGCCTCAAGATAAGGA 98958
       |||||
Qy   1730 TGATGTCGAGAATGGGCGCGGGGGGGGATTCGGGGGGTGAGGGCTGAAAATCAGCCTGC 1789
       |||||||
Db 98959 TGATGCTCGAGAATGGGGTGGGGGAGAGAC-------- 98987
       |||
Qy   1790 GCCTGGAACCAGAGAAATCTGGAAGGCTGAAACGATACCATAAAGGGAACCCAGGCT 1849
       |||||||||||||||
Db 98988 ---GGGACTGAGACTCTGGAAGCGCCTGGAAACGATACTGTAAAGGGGAAAACCCAGGCT 99043
       ||||||
Qy   1850 ACCTTTAGATGTAAATCATGAAAGACAGGAAACAGGGAGAGACGGGAGAGGACC 1909
       |||||||||||||||
Db 99044 ACCTTTAGATGATGAAGTCATGAAAGTCAGGGAGCAGGGGAGCTGAGAGAGTAGAGG---- 99099
       |||
Qy   1910 CCGGGGCAAGACATGGAAGCAAGGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGC 1969
       ||||||
Db 99100 ----AATAAGACAAGGACAAGGACAAGCCAGGTTGAGTGAGTCGTTGTGAT- 99139
       |||
Qy   1970 TGAAGGCAGAGCCCTGGTATGAGCACCAGAACACAGCCTAGGTTAATGTCGAGACA 2029
       |||||
Db 99140 --------TCCGATATGAGCCCGGAACAGCAGCAGCACAGAGGCCGATTTAGAGACA 99187
       |||||||||||
Qy   2030 GGGAACAGAGGTAGACAGAGGAACAGCAGAGACAGAGACAGCGGGGGGCAGCCAGGAGGACC 2089
       |||||||||||||||||
Db 99188 GGAACTGAAGGTGGCACAGCAGGCAGCGGGGATCCAGGTAACAAGGGATG 99247
       |||
Qy   2090 GTCCTTCTCACCTGTCGCCAGCGTCCATCT-GTGTCCACATACTCTAGAATGTTCATC 2148
       || |||
Db 99248 GTGCCCTCACCTGTCGCCAGAGCGTCCATCTGGGGTCCACTTACTTCCTCTAGAACGTTCATC 99307
       |||||
Qy   2149 AGACTGCAGGGCTGCTTGGGAGGCGCCTGGTTTGGGAGGAGACTA 2208

       ||||||||||||||||||||||

Db 99308 AGACTGCAGGGCTGCTTGGGAGGGCGCTGAGAGTATGTGAGAG--CGGGGGAGACTA 99366

```
Qy    2209  GGGGGCCTAGGAGAAGGAAGAGGGCAAACCAGGCCACACAAGAGGGGCAGAGCCCAGAA  2258
            |||| ||| |||||||||||||| ||| ||||||||||||| |||||||||||||||||
Db   99367  GGGGACCGAGGAAGGAAGAGGGAGAGAGGGGCAAACCAGGCCAGAGGGGCAGAGCCCAGA  99426

Qy    2269  CTGAGTTAACTCCTCCTTCCTTGT--ATCTCCATAGGAGGCAGTGGCAATCTAGGTTCC  2328
            |||| |||||||||||||  ||||  ||||||||||||||||| |||||| | || |||
Db   99427  CTGATTTAACTCCTCCTTGT---ATCTTCCATAGGAGGCAATGGAAACTCTGTGCCCAC  99483

Qy    2329  CATCCCCATGAGCCCCACTACCACCAAGTTGGCCTGAGTTCTAGGTTCC           2388
            |||       |||||||||| ||||||||| |||||||| |||| || |
Db   99484  CAT--------CCCCACTACCTACCAAGTTTGGCCTGAGTAGCATAGGTTCC         99533

Qy    2389  CTGAGACAGAGCCTGGCCTTTGTCTCTTGGACCTGACCCAAGCTGACCCAATGTTCTCA  2448
            |||||||||| ||||||||| ||| ||||||||||||||||| ||||||||||||| |||
Db   99534  CTGAGACAAAGCCTGGCCTTTGTCTCTTGGACCTGACCTGACCCAATGTCTCA        99593

Qy    2449  GTACCTTGAATGCCCTCAAGAGCTTGAGAATCCAGGCAGTGACATATTAGGCCATGGGCT  2508
            ||| |||||||| |||||||||||||||| ||||||||||||||||||||||| |||||
Db   99594  GTGCCTTATCATGCCCTCAGAGCTGAGAATCCAGGCAGTGACATATTAGGCCATGGGCT  99653

Qy    2509  AACCCTGGAGCTTGCACACAGGAGGCCTCAGGGACCAGCTGCAGACAGG            2568
            ||||||||| |||||||||||||||||||||||||||| |||||||||||
Db   99654  AACCTTGCAGCTTGCACACAGGAGGCCTCAGGGACCAGCTGCAGACAGG           99713

Qy    2569  TGGCCTTTATCCCAAAGACAACCATTTGGCATAGGTGGCTGCAAATGGAAATGCAAGG   2628
            ||||||| |||  ||||||||||||||||||| ||||||| | || ||||| |||||||
Db   99714  TGGCCTTCATCCTCGGAGCACAACCATTTGGCAGCAGATGGGAACGCAAGG         99773

Qy    2629  TTGAATCAGTCCCTCAAGAATACTGCATGCAAGACCTAAGACCCCTGGAGGAGAGGGGT  2688
            ||||| ||||||||||||||||| |||||| |||||||||||||||||||| |||||| |
Db   99774  TTGAAACAGGTCCCTCAAGAATACCGGCGTACTCAAGAGCTAAGACCCCTGGAGGACTGT 99833

Qy    2689  ATGCTCCTGCCCCCACCCACCATTAAGGGAGTGAACTATCCTCAGGGGCTGGCGACCTTG  2748
            |||||||||||||||  |||||| |||||| |||||| ||||||||| || || ||||||
Db   99834  ATGCTCCTGCCCCCATCCCCATTGAGGTATTGAACTATCCTCAGGGGCTGGCAACCTTG  99893

Qy    2749  GGGAGACCACCACTACTGAGAGTGCTGAGCCAGGAAAAACTGACCGCCCTGTGTCCTGC  2808
            ||||||||||||| |||||| |||||||||||| |||||| |||||||| || |||| |
Db   99894  GAGAGACCACCATCACCGAGAGTGCTGAGCCAGGTGTGTCACCGCTGTGTCCTAC    99953

Qy    2809  CCACTCCACACTCTAGAGCTATATT--GAGAGTGACAGTAGAGGTGGGAGCTGGT    2866
            ||||||||||||| ||||||| |||  ||||| |||| || ||||||||||||||
Db   99954  CCACTCCACACTCTAGAGCTATATTGAGAGAGTGACAGTAGAGGT-GGGGGAGAGCTGGG 100012

Qy    2867  AGCAGGGAGAGTGTTCCTGGGTGTAGGGGTAGGGGAAAGCAGGGGAGGTCTCG      2926
            |||||||||||||| |||||||| ||||||||||||||||| |||||||||||
Db  100013  AGCAGGGAGAGTGTCCTGGGTCGTCGGGGGCGTCGGGGGAGTCAGGGGAGGGGAGTCTGG 100071

Qy    2927  CTTTGTCTCTGAACACAATGTCTAGTTGTAATAAAAA                      2986
            ||||| ||||||||| || |||||  |||| ||||||
Db  100072  CTTTGTCTCTGAACACAGTGTCTACTTAGTTATAAACATGCATCAGCCCTAAAGACCCA 100131

Qy    2987  ACATCTACAGACCTCTGAAAGAAAACAGCAGGGGTTGTCTCTGAGCCTTG        3046
            ||  ||| || ||| |||||||   || || || |||||| |||||||||
Db  100132  ACACCTACACCTGAAAAAACAGCAGGGCCCGGTTGCTGCCGTTGGGCCTTG       100191

Qy    3047  GGTGCTTGATGGTGCCACAAAGGAGGCATGAGTTGAGTAAGGCCGTTA          3106
            ||||||||| |||||||||| |||||||||||| ||| |||||| |||
Db  100192  GGTACTTGATGGTGCCACAAAGGGGCATGTGAGTAGTATAAGGCCCCAGGGGTTA    100251

Qy    3107  GAGAAGGGACACTTGGGACGGGGCCTCAGTGGGTCAGTCTGCAGAGCCT          3166
            |||| ||| |||||||||| |||||| |||||||||||| | ||||||||
Db  100252  GAGAAGGGACACTTGGGACGGGGCCTCAGTGGGTCAGTCTGCAGAGCCT         100311

Qy    3167  GGGGCCAACTGGTGTAAATCTCTGGGCATCAAAGCACCTGCATCCTC           3226
            ||| ||||| ||||||||||||||||||||||| || ||||||||| |
Db  100312  GGGGCCAACTGGTGTAAATCTCTGGGCCTGCCCGAGTAGCACCTGCATCCTC      100371

Qy    3227  TGGCAGCCTCGGGAGCGAGGGCGAGAAGAGGCGAGGGGAGACCTTTCTCCCCTCAGCCC  3286
            ||||||||| |||||||||||| |||||||| ||||||||||| || |||||||| | ||
Db  100372  TGGCAGCCTCGGGAGGGGGAAGAGCGGAGAAGGGAAGACCTTTACCACTCCCATCAGCAC 100431
```

```
Qy    3287  CAGGATTAACACCTCTGGCCTTCCCCTTCCCACCTTCCCATCAGGAGGTGGGGTTGCAG  3346
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db  100432  CAGGATTAACACCTCTGGCCTTCCCCTTCCCATCCCATCAGGAGTGAGGGTTGCAG     100491

Qy    3347  AGGGAGGGTAAAACCTACACATGTCCAAACATCATGTGCACGATATGGCACTAGTATGT  3406
            |||| |||||| | | || |||  ||   ||  |  |||||| || ||||||||||| ||
Db  100492  AGGGAGGGTAAGACTAGAT--------AATTGTGTGCATGATATATGACTAGTATGT     100543

Qy    3407  GTAGAGGCAAGAAACGAAATCTGCA------GGCTAACTGGGTTAAATGTGTAAA---GT 3457
            | ||||||||||| |||| |||||||      |||||||||||||||||| |||
Db  100544  GTAGAGCAAGAAAGGAAATCTGGAGTGAAGTGTGTAAAGTCTCTGTGTGTGTGCATGTGC 100603

Qy    3458  CTGTGTCATGTGTGTGTCTGACTGAGGGCATGGGCATGAGTTGTGTTGTGTGCAGCTGTT  3517
            ||||||||||||| | ||||||||| |||||||||| |||||| ||| |||||| |||||
Db  100604  ATGTGAATGCATGTGTGTGTCTGACTGAGGGCATGAGTTGTGTTGTGTGCAGCTGTT    100663

Qy    3518  GTGCGTGAGGTTACCAGACTGCAGGTTTGTGTGTGTAAATTGCCCAAGGCACAGTGGGTGAA  3577
            ||| |||  ||| |||||||| |||| |||| ||||
Db  100664  GTATGTGACATTACCAGACTGCAGGTTTGTGT----------AAGTCAGTGAG        100708

Qy    3578  TCCCTTCCATGGTTAAAGAGATTGGATGATGGCCTGCATCTCAAGGACCATGGAAAATA  3637
            |||| ||||||||||| | | || | ||| || |||| || |||| |||| || |||||
Db  100709  ACCCTCCCATGGTTAAAGAAATTGGATGATAGCCTGCATCGGGGCAACATTGAAAGTA   100768

Qy    3638  GAATGCACACTCTATATGTGTCTAAGCTAAGGTGACAGTATCTGTACAGACACCTGT    3697
            ||||| ||||| || ||||| ||||| ||||| |||||||| ||| ||||||||  ||
Db  100769  GAATGGACACTCTTTATTATATGTGTCTCTAAGCTAAGGTCAAGTAGCAAGTCTTTGGAC 100828

Qy    3698  CTAGAGATGTGGCAACAGACTACAGACGTATCGTGTACAGAGTAAGGTAAGAGAGAGA   3757
            |||| |||||||| ||| ||  | |||| ||| | ||||| || ||  |||| | |||
Db  100829  CTAGAGATATGGGCAACAGACAATAGACAGTCTAAAGTACAGAGTAA---AGAGAGAGG  100886

Qy    3758  GGGGTGTAGAATTCTTACTATCAAAGGGAAAGTGAGTCGTGCACCTGCAAAGTTGGAT   3817
            ||||||| ||||| ||| | |||||| ||| |||||| || |||||||| || |||||
Db  100887  TGAGGTATAGAATTCTCTTATTCTCTATTGAGAGGGAAAATGAATGACCTGCAAAATGAAT 100946

Qy    3818  GCTCTCCCTAGACATCATGACTTTGTCTCTGGGAGCCACTGTGCAACTTCAGGTCT    3877
            ||||||||| | |||| ||| |||| | |  ||||||| |||| | ||| |||||||
Db  100947  GCCCTCCCTAGACATCATGACTTTGTGTGTCCCTGGGGGGCCAGCCACTGTGGAACTTCAGGTCT 101006

Qy    3878  GGGAGACTAGGAGGCTCCCCTCAGCCTGCAGTAGAGTCAGATAGGC----GGG       3935
            ||||||||||||| | || ||||| ||| ||||||| |||||||||    |||
Db  101007  --GAGGTAGGAGGCGCCCCTCAGCCTGCAGTAGTCGCAGATAGCTAGGGTGTAAAGAGG 101064

Qy    3936  GGAGGGGGAGAGCCTGGGGATGGGAAGCCAGAGAGGACAGAGCTGGG            3995
            |||| ||| ||| |||| || |||| ||| || || |||||| |||||
Db  101065  GGAAGGGGAGAGCCTGGGGATGGGAAGCCAGGAACAGGGGGACAAAATATTAGGCCC   101124

Qy    3996  GGAAGAAGCTGACCCTACCCAGT-TGTTCAACTCACCCTTCAGATTAAAAATAACTG    4054
            |||||||||| ||||| ||||||  |||||||||  | |||| | ||||||| ||| |
Db  101125  GAAGAGACAGTGACCTTGGGTGTGTTGTTGAGACGTGTGAGACAGCCTTCAACTAACTA 101184

Qy    4055  AGGTAAGGGCC---TGGGTAGGGGAGGAGAGAGGAGGCGCCCTGCTCCTGTCTCCTCATG  4111
            |||| ||||||   ||||| |||| |||||| || |||| |||||| |  | ||||| ||
Db  101185  AGGTAAGGGCCATGTGGGTAGGGGAGGAGAGACGGTGTGAGAGGTTCCGTGTCCTCTATCTG 101244

Qy    4112  CCCTGAGGCCCTTTGGGGAGGAACAACAGACCTTTCAGGAATTCGGAGGCGTTA      4171
            ||| || ||||| |||||| || | |||||| || ||  |||||| ||||||||
Db  101245  CCCATCGGCCCTTTGGGGAGGAATGTGCCCCAAGGGGGATGTGCCCCAGGGGATTA  101304

Qy    4172  GAGGGGCGAGGGCGACACAGACCTTTCATGGCAAGTTGGGCCCGTAGTGATCGATTGA    4231
            ||||| |||| || | | ||| ||| |||| | ||||||||| |
Db  101305  GAGGGCGAGGGCAGCAGCAGACTTTCATGGCAAACTTCAGGGC-------------   101347

Qy    4232  CAAGAACTCGCCAATCGACTACCCTTCTTCTTCTAACGGACAGGAGGAGGAACTCGAGTTTAC  4291
            ||||||| ||||| ||  |||| |  || |||| |||| |||||||||||||||||||||| |
Db  101348  ------------TGCTGTCCTCCTCGTGTCACCTCCAGAGCCCAAGGGATCAAAGGAGGAG 101392

Qy    4292  CACTCCCATCAGTGATAGAGAAAAGTCGAAAGTCGAGTTTACCACTCCCTATCCAGTCGATA  4351
            |||||| || ||||||| | |  |||| |
Db  101393  GAGCCAGGACAGAGGAGGGGATGGATGGGAGGGAG----------------------     101420

Qy    4352  GAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCCAGTCGATAGAGAAAAGTCGAAAGTCGA  4411
```

```
Db  102123  TGAGAATGCCCCAGTTTCCCTGTAGACAGCAGATCATGATTTCCCAGAAGCCAGAC  102182
                                                                     
Db  101421  ---------------------------------------------------------  101420

Qy    5485  TTCCAGCCCTCCTTTCTGCCCAGTCGCCCCGGCACCTCTTAGCAAACCTCAGCCACCCT   5544
                                                                     
Qy    4412  GTTTACCACTCCCTATCAGTGATAGAGAAAGTGAGTTACCACTCCCTATCA          4471

Db  102183  TTCCAGCACCCGC---CCTCTGCCCAGCAACCTGCCACTCTTAGCAAACTTCAGCCACCCT  102240
                                                                     
Db  101421  ---------------------------------------------------------  101420

Qy    5545  TACCCCACATAGACC-------------                                  5559
                                                                     
Qy    4472  GTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAA    4531

Db  102241  TCCCCCACATAGACCAAGTCTTGCAGAGAGCCTCCTTCAGATGACTTCGAGTTTCTGCAA  102300
                                                                     
Db  101421  ---------------------------------------------------------  101420

Qy    5560  ----------------TCTGACAGAGAAGCAGGCACTTTACATGAGTCCTGGTGGGAGA    5603
                                                                     
Qy    4532  AGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGCTCGGTACCAG    4591

Db  102301  GGAAAGGAGAACTCTTTGTGATGGGAAGCAGGCAGGCACTTTACACGGAGTCCTGACGGGAGA  102360
                                                                     
Db  101421  ------------------------------GGTCCCAG                       101428

Qy    5604  GCCATAGGCTACGTGTAAAAGAGGCAGGGAAGTGGTGTGTAGGAAAGTCAGGACTTCA    5663
                                                                     
Qy    4592  CAGAGGACTCCAAATTAGGCAGCAGGCATATGGGGATATAAAGGGGCTGGAGCAC       4651

Db  102361  GCCATAGGCTATGGCATAGCAGCAGGCAGGGAGGTGGTGAAT---------TGGACTTCA  102411
                                                                     
Db  101429  CAGATGACTCCAAATTTAGGCAGCAGGCACGTGGAATGAGCGTATAAAGGGGCTGGAGCGGC  101488

Qy    5664  CATAGAAGCTAGCCCACACAGAAATGACAGACAGATCCTCCTATCTCCCCATAAGA       5723
                                                                     
Qy    4652  TGAGAGCTGTCA----GAGATTTCTCAACCCAGGTAAGAGGGAGTTTCGGGTTGGGGGCT   4707

Db  102412  CACAGAAG-CTAGCACACACACCAGGAAT------GACATATCCTCCTATCTCCCCATAAGA  102466
                                                                     
Db  101489  TGAGAGCTGTGACCGGAGATTTCTCCATCCCAGGTAAGAGGAGTTTAGCGTGGGGGGCT  101548

Qy    5724  GTTTGAG    5730
                                                                     
Qy    4708  CTTCACCCACCACAGACCTCTCCCCACCTAGGAGGAAACTGCCTTCCTTCCTGGAGTTGGGT    4767

Db  102467  GTTTAAG  102473
                                                                     
Db  101549  CTCCAACCGCACCAGACCTGTCCCACCTAGAGGGAGAGTGTCTTCCCTGGAAGTTCTCAGTGGC  101608


RESULT 5
AC119293/c
LOCUS       AC119293          210784 bp   DNA     linear   HTG 19-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-272D8, WORKING DRAFT SEQUENCE.
ACCESSION   AC119293
VERSION     AC119293.4  GI:25072533
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 210784)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,A., Baca,S., Amin,A., Anguiano,D.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,K., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,A., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hernandez,J.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lope,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., McNeill,T.Z., Meenen,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
```

TITLE   Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
          Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
          Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
          Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
          Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
          Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
          Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
          Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
          Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
          Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
          Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F.,
          Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
          Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
          Yu,F., Zhang,J., Zhao,J., Zhou,X., Zhou,S., Dunn,D., von
          Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
          Weinstock,G. and Gibbs,R.A.
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 210784)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (26-APR-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
REFERENCE 3 (bases 1 to 210784)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 19, 2002 this sequence version replaced gi:23602929.
          The sequence in this assembly is a combination of BAC based reads
          and whole genome shotgun sequencing reads assembled using Atlas
          (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
          in the feature table below represents a scaffold (a 'contig-scaffold'
          assembly (a 'contig-scaffold'). Within each contig-scaffold,
          individual sequence contigs are ordered and oriented, and separated
          by sized gaps filled with Ns to the estimated size. The sequence
          may extend beyond the ends of the clone and there may be sequence
          contigs within a contig-scaffold that consist entirely of whole
          genome shotgun sequence reads. Both end sequences and whole genome
          shotgun sequence only contigs will be indicated in the feature
          table.
          ------------------ Genome Center
             Center: Baylor College of Medicine
             Center code: BCM
             Web site: http://www.hgsc.bcm.tmc.edu/
             Contact: hgsc-help@bcm.tmc.edu
          ------------------ Project Information
             Center project name: GRMH
             Center clone name: CH230-272D8
          ------------------ Summary Statistics
             Assembly program: Phrap; version 0.990329
             Consensus quality: 180229 bases at least Q40
             Consensus quality: 182328 bases at least Q30
             Consensus quality: 183274 bases at least Q20
             Estimated insert size: 184593; sum-of-contigs estimation
             Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
          ----
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank-draft-data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          *   consists of 1 contigs. Gaps between the contigs
          *   are represented as runs of N. The order of the pieces
          *   is believed to be correct as given, however the sizes
          *   of the gaps between them are based on estimates that have
          *   provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *   210784: contig of 210784 bp in length.
FEATURES   Location/Qualifiers
     source   1..210784

     misc_feature   1..1742
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-272D8"
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature   18119..19359
                     /note="wgs_end_extension
                     clone_end:Sp6"
     misc_feature   21232..22032
                     /note="clone_boundary
                     clone_end:Sp6"
     misc_feature   200775..201644
                     end_sequence:BZ243354"
                     /note="clone_boundary
                     clone_end:T7"
     misc_feature   209704..210784
                     site:
                     end_sequence:BZ243352"
                     /note="wgs_end_extension
                     clone_end:T7"

ORIGIN

Query Match   40.9%; Score 2343; DB 2; Length 210784;
Best Local Similarity   71.2%; Conservative 0; Pred. No. 0;
Matches 4321;   Mismatches 795; Indels 951; Gaps 55;

```
Qy       1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCACCAGGTGCCCAGTCTCAATTTCAGT 60
           ||||| |||||||||||||||||||||||||||||||||||||| |||||||||||||
Db  153822 GGATCCTCAAGGTCACACAAGGGTCTCCACCTACCAGGTGCCCAGTCTCAACTTCAGT 153763

Qy      61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCAGAGCTAATTTGGACTTTGTTTTTAT 120
           ||||||||||||||||||||||  |||||||||||||||||||||| |||||||||||
Db  153762 TTCCATGTGCTTGTCCCTACAATGTTGGCCTCCCAGAGCTAATTTTGAC-TTGGTTTTAT 153704

Qy     121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTCGTACCCAGTCTCAAGGGTGCCGTGA 180
           ||||||||||||||||||||||||||||| |||||||||||||||||||| ||||||||
Db  153703 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTCTACCCAACCCTAAGGGATGCCTGTGA 153644

Qy     181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 240
           |||| ||||||||||||||||| |||| ||||||||||||||||||||||||||||||||
Db  153643 AGCCCTGAGACCTGGAGCCTTTGAAACAGCACCTTAGGCAGAAACAACAATAAAGCAATTT 153584

Qy     241 TCCTTAAAGCCAAAATCCTGGCCTAGACTCTTCTTCTCACTGACCTCGGTCCT-GGGCTC 299
           ||||||||||||||||||||||||||||||||||||| |||||||||||| |||||||||
Db  153583 TCCTTCAAGCCAAAATCCTGGCCTCTCTAGACTCTTCTTCTCACTGACCTCCGGTCCTGGGGCTC 153524

Qy     300 TAGGGTGGGGAGGTGGGGCCTTGGAAGAAGAAGGTGGGAAGGTCAAAAAGCCGATCCCTA 359
           ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db  153523 TAGGGTGGGGAGGTGGGGACTTG---AAAGAAGGAGGGAGGTGGGGAAGGGCCACTCCCCA 153467

Qy     360 GGGCCCTGAAGTTCGGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCTCCAGC 419
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  153466 GGACTCTGTGAAGTTCTGAGCCTTCCCAGTACAGTGCTCGGCTCATAGATCCTCCTCCAGC 153407

Qy     420 CAAACATAGCAAGAAGTGATACCTCCTTTGTGACTTCCCCAGTACCTGTCAGGT 479
           |||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db  153406 CAAACACAGCAAGAAGTGATGCTTCCTTGTGACTTCCCCAGTGCCCTGTCGTCCGGT 153347

Qy     480 TGAAACAGGATTTAGAGAGCCTCGAACTCACCTGAAGCTCATCCACCAAGC 539
           ||||||||||||||||||||||||||||| ||||||| ||||||||||||
Db  153346 TGAAACAGGATTTAGAGAGCCTCTAAACTCACCTGGAGCTCTGAGCCCATCTAC----C 153291

Qy     540 AAGCACCTAGGTGCCACTGCTAGTTAGTATCCTCACGCTGATAATATGCGAGAGCTG--GGC 597
           |||||||||||| |||||||||| ||||||| |||||||||||||||||||||||||
Db  153290 ACACACCTACATGCCCGGTTCTAATCAGTCGCCTCAATCAGTATCCTACACTGATAATAATACGCGGCAAAGGTGACCAC 153231

Qy     598 CACAGAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCTGCAAAGGTATGACC 657
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  153230 GGGAGAGGGCCTGGAGTGTAGGAACTGGCCCAGTGACTTTTCAGTCTGCAGAGGTATGACC 153171
```

```
Qy      658 CCCTCAGCAGAT---GTAGTAATGTCCCCTTAGATCCATCCCAGGCAGGTCTCTAAGAG    714
            |||| || |||||   ||| ||| ||||||||||||||| |||||||||||||||| |||||
Db   153170 CCCTTGGCCAGATATGGGCAGATGTGGCAGATGTCCCCTTAGAACCCATCCCAGGCAGGTCTCTGAGAG 153111

Qy      715 GACATGGGATGAGAGATGTAGTCATGTGCATTCCAAACACAGCTATCCACAGTGTCCT    774
            ||||| |||||||| ||||| |||||||||||||||||||||| |||||||||||||||
Db   153110 GACACAGGATGACAGAGATGAGTCATGTGCATTCCAAACATCCAAACACAGCTATCCACAGTGACTCT 153051

Qy      775 TGCCCCTTCCACTTGACCAGGAGAGCAGTAACCTTAGCCTATCTTCTTCCTCCCCATCC    834
            || |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   153050 TGTCCCTTCCACTTGACCAGGAGAGCAGTAACCTTAGCCTATCTTCTTTCTTCCTCCCCATCT 152991

Qy      835 TCCCAGGACACACCCCCTCGGTCTCCAGTATTCATTCTTCCTCACGTCCCCTGTGAC    894
            |||||||| ||| || ||||||||| ||||||||||||| |||  |||||||| |||
Db   152990 TCCCAGGACCCCAACCCCCGGTCTCATGGTATTCATTCTT-CTTCATTTCCCCCATCCAC 152932

Qy      895 TTCCATTTGCAAGGCTTTTGACCTCTGCAGCTGCGTGCAGTAGAGTTTGGCCTGGCTG    954
            ||||||| |||||||||| |||||||||||||| ||||||||||||||||||||||||||
Db   152931 TTCCATTTGCAGGGCTCTTTACCTCTGCAGCTAGCTGCGTGCAGTAGAGATGGAGCTCTATATG 152872

Qy      955 TGGCAAGCCATCTCAAGAGAAAAGCACAGCAGGGGGGCCTGGGTAGAAAAAGGATCAGGA 1014
            || |||||||||||||||||||| || |||||||| || |||||||||||||
Db   152871 TGGCTCTGAAAGGACAATGGAAGACATCTCTTAGTCTCCCCTCAACCCCACCT 152813

Qy     1015 ACTAAATCACTGGGGGCCTGGGGGTAGAAAAGAGTGAGTCCGCTCCAGCTAAG 1074
            ||||||||||||| | ||||| || || ||| |||
Db   152812 ACTTTAGATGTAAATCATGAAAGAGAAGGAGAGGAGAGCTGAGAGGAGGAGGGAC 152761

Qy     1075 CCAAGCTAGTCCCCGAGATACTCTGCCACAGCTGGGTGCTGCTCGGGTAGCTTTAGGAATG 1134
            ||  || |||||| ||| |||||| |||||||
Db   152760 CAAAGCCACCGAATCCTTGGGTATT----CACAGCTGGGTGCT-GAGGTAGGTTTTAGGAATG 152706

Qy     1135 TGGGTCTGAAAGACAATGGAGAACATCTCTTTGAGTCTCCCCTCAACCCCCACCT 1194
            |||||||||||||| ||||| |||||||||||||||||||| |||| |||| ||
Db   152705 TGGCTCTGAAAGAACATGGGATTTGAAAACATCTCTTTGAACCTGCCCGCTCTGGGCT 152646

Qy     1195 ACAGACACACTCCCCTCCCCAACCCCAGAC 1254
            |||||||| ||||||| || |
Db   152645 AAATCTCCCCCTCCAACCCCATGGCCAGACTG 152596

Qy     1255 AGCTAGGCAACGAACAGAGACATGGGCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAA 1314
            ||||| |||| || ||||| ||||||||| ||
Db   152595 AGTTAGGCAACTGGAGACAATGGGCCCTGTG-TGAGGACGAGG----GGTGCCAATAGCAA 152541

Qy     1315 AAACAGCAGGGGGGGGGGACAGAGAACAGAACAGAGGAAGGAAAAGGCCAGTCA 1374
            ||||||| ||||| || || |||||| |||||| |||| |||| || |||
Db   152540 AAACAGCAGGTGGGGACAGGGATTTGAAAGGAGGAACGAACCGAACCGGCC---CA 152484

Qy     1375 ATCAGATGCAGAAGATCAGAATGGGAAGCCAACAACAATTTTCTAGTT 1434
            ||||||| |||||| ||||||| ||||| ||||||| |||| ||||
Db   152483 ATCAGATGCAGAAGATCAGAGATGGGAAGCCAACCAATGCATAGCTTGAACAGAAAATAGAAG 152424

Qy     1435 AGGGAGAGA 1443
            ||||||||||
Db   152423 AGGAAGAGATGGCCAGAATCATTAGGGTGGACTTCCTCCATTAAGATGCCTCTGTGT 152364

Qy     1444 -------------TTCTGGGCATAAGGAGGCCACAGAAAGAGAGCCCAGGCCC 1484
                         || |||||||
Db   152363 CCGACTCTACATCCCTGGGTACTTGGTCATTGGCTCAGCTCAGGAATAGCAGACACAGAGAACT 152304

Qy     1485 CCCAAGTCTCCTCTTTTATACCCTCATCCC 1513
            ||||||| || |||| |  ||
Db   152303 GTATATGGCCATTTTCACATTCTCAAATCCTGAAGCTTGACAACTTGACAATTTCTAGT 152244

Qy     1513 -------------GTATATTGTCTTCTATACAGATATGTCTTCATGGACAGTGCTTCCCTGAAT 152184
Db   152243 TACATTGTGTTTACTCTTATCTATACAGATATGTCTTCATGGACAGTGCTTCCCTGAAT 152184

Qy     1513 ------------------------ 1513
Db   152183 ACAAGGTCCTTTTTTTATGAAATCAGTGTGCATTAAATCAAGTCTGCATAGGGAGGCCA 152124

Qy     1514 -------------GTCTCCCAATTAAGCCCACTCTTCTTCCTCAGATCAGAC 1551
Db   152123 ...


Qy   152123 CAGGAAGAGCCCAGGTCCCAGGTCTCCTCTTTATGCCCTCCTAAGCCCATTC 152064
            ||| |||||||||||||| | |||||||||||| | || ||||| |||||||||
Db     1552 CTGAGCTGCAGCGAAGAGACAGACCCTAGGGAGGAGATCACACTGGATGAAGAGAGATGTGGAG 1611

Qy   152063 CTCCTC-CTGGAACAGACCTGAGCGACCTGAGGTCTTAAAGAGGATCA 152005
            |||||  || |||||||||||||| ||||||||||| |||| | ||
Db     1612 AGTCCGTAGAGTCTTAAAGAGGATCAAGTCCAGGCAACCTAAGAGCCAGAGAGAGAAAGTGCTTCAAAG 1671

Qy   152004 CAGTGGAAGGAGATGTGCACCAACACAGCCTCAAGAACAAAGCAGCAAAG 151945
            ||||||||||| |||||| |||||| ||||||||||||||||
Db     1672 GTGGCCAGCTGTG--CACCACAGGGTCGAGGACTGGTGGTAGAGCCTCAAGATAAGGA 1729

Qy   151944 GCAGCTAGCCCGTGCACAGACTACTAGCCGTCAAGAGTAAGGA 151885
            || ||||| ||||  ||||||| ||| ||||||||||| ||||||
Db     1730 TGATGCTCAGAATGGGCGGGGGGGGATTCTGGGGGGGAGAGAGGTGAGAGGA 1789

Qy   151884 TGATGCTCAGAATGGGGTGGGGTGAGGAGAGAC 151856
            |||||||||||||||||| |||| ||||||||||
Db     1790 GCCTGGAACAGAGAATCTGAAGGCTGGAAAACCATACCATAAGGGAGAACCCAGGCT 1849

Qy   151855 ---GGGACTGAGACTCTGCAGTAGCTGAGAGGTGAAAACCCAGGCT 151800
               ||||||||||  || ||| | ||||| |||||||||||||
Db     1850 ACCTTTAGATGTAAATCATGAAAGAGAAGGAGAGGAGAGCTGAGAGGGACC 1909

Qy   151799 ACCTTTAGATGAAGTCATGAAAGTCAGGGAGAAGGAGAAGGAGAGAGG------ 151744
            ||||||||||| |||||||||||  | || ||| |||||||
Db     1910 CCGGGGCAAGACAATGGAAGCAGACAAGGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGC 1969

Qy   151743 -------AATAAGACAAGGACAAGCAAGCCAGGTTGAGTGCTTGCTCTGTGAT------ 151704
                   |||||| |||||| ||||||||||||| |||||
Db     1970 TGAAGGCAGAGCCCTGGTATGAGCACAGCAGAACAGCAGCAGCAGCACAGATGGGTTAATGTCGAGACA 2029

Qy   151703 -------------TCCGATATGAGCCGTGAGGAACAGAGGAACAGAGGAAGAGACA 151656
                         ||||| ||| |||||| |||| |||||| ||||||
Db     2030 GGGAACAGAGGAGTAGAACACAGAACAGAGGAAGCAGCGGGGGGAGCCCAGGAGTAACAAAGGAATG 2089

Qy   151655 GGGAACTGAAGGTGGCACGACAGGGGGAGCCCAGGAGTAACAAGGGATG 151596
            |||||| |||||| || ||||| | |||||||||||||||||| ||||||
Db     2090 GTCCTTCTCACCTGCCCAGGATGGCCAGGTGTCACATACTCTAGAATGTTCATC 2148

Qy   151595 GTGCCTCTCACCTGGCCAGCGCTGGGTGTCCAGTTGGTTCCACTTACTCTAGAATGTTCATC 151536
            |||| |||||| ||||||| ||  | ||||||| ||| |||||||||||||||||||||||
Db     2149 AGACTGCAGGGGCTGGCTTGGGGAAAGAGTATGTGAGAGCCAGGGGAGACAA 2208

Qy   151535 AGACTGCAGGGGCTGGCTTGGGGCTGGTTTGGGAGGGGCTGAGAGTGTAGAG--CGGGGGGACTA 151477
            |||||||||||||||||| ||||  ||| |||  || ||||| ||||
Db     2208 GGGGCTGGCCTGGGGAGAAAGGAGGCAAACCACCAGGCAAACCAGCCCAGAAC 2268

Qy   151476 GGGGGGCCTAGGAGCAGAGATGGGAAGGAGGCAAACCACCAGGCAAACCGGCC---CA 151417
            |||| ||||| | ||||| ||||| ||||| ||||| |||||| |||| ||  ||
Db     2269 CTGAGTTAACTCCTTCCTTCT---ATCTTCCATAGGAGGCAATGGAAATCTCGTGTGACCAC 2328

Qy   151416 CTGATTTAACTCCTTCCTTCCTTGT---ATCTTCCATAGGAGGCAATGGAAAATCTCGTGTGCCCAC 151360
            |||| ||||||||||| |||  ||     ||||||||||||||||||||||| | |||||| ||||
Db     2329 CATCCCCATGAGCCCCACCTACCAGTTTGGCCTGGCATTCTAGGTTCC 2388

Qy   151359 CCCCCCACCTACCAGTTTGGCCTGGCATTCTAGGTTCC 151310
            |||||||||||||||||||||||||||||||||||||
Db     2389 CTGAGGACAGAGGCTTGAATGCCTCAAGCTGACCCAATGTTCTCA 2448

Qy   151309 CTGAGGACAAAGCCTGGTTTTTGTCGGTTGGACCTGACCCAATGTTCTCA 151250
            |||||||| ||||||||| |||||| ||||||||||||||||||||||||||
Db     2449 GTACCTTTGAATGCCTTCAAGAGCTTGAGAAGCACAGGCCATGGGGCT 2508

Qy   151249 GTGCCTTATCATGCCCTTCAAGAGCTTGAGAAGCACAGGCCATGGGGCT 151190
            || |||| ||||||||||||||||||||||||||||||||||||||
Db     2509 AACCCTGAGCTTGCACACAGACTGCCTCAGTGCCAGACAGG 2568

Qy   151189 AACCTTGCAGCTTGCACACAGACTGCCTCAGTGCCAGACACAGG 151130
            |||| |||| |||||||||||||||||||||||||||| |||||
Db     2569 TGGCCTTTATCCCCAAAGAGCAACCATTTGCATATGGGAATGGGCTGCGAAATGGGAATGGCAAGG 2628
```

This page contains a two-part DNA sequence alignment between a query (Qy) and a database (Db) sequence. The alignment is presented in paired rows (Qy over Db) with flanking position numbers. The dense nucleotide sequence content is not legibly resolvable for faithful base-by-base transcription.

Right-hand alignment block, row labels and position numbers:

Row	Start	End
Db	151129	151070
Qy	3698	3757
Db	150014	149957
Qy	3758	3817
Db	149956	149897
Qy	3818	3877
Db	149896	149837
Qy	3878	3935
Db	149836	149779
Qy	3936	3995
Db	149778	149719
Qy	3996	4054
Db	149718	149659
Qy	4055	4111
Db	149658	149599
Qy	4112	4171
Db	149598	149539
Qy	4172	4231
Db	149538	149496
Qy	4232	4291
Db	149495	149451
Qy	4292	4351
Db	149450	149423
Qy	4352	4411
Db	149422	149423
Qy	4412	4471
Db	149422	149423
Qy	4472	4531
Db	149422	149423
Qy	4532	4591
Db	149422	149415
Qy	4592	4651
Db	149414	149355
Qy	4652	4707
Db	149354	149295
Qy	4708	4767
Db	149294	149235

RESULT 6
AY191158
LOCUS   AY191158   5487 bp   DNA   linear   ROD 22-SEP-2003
DEFINITION   Rattus norvegicus beta myosin heavy chain gene, partial cds;
                beta-alpha myosin heavy chain intergenic spacer region, complete
                sequence; and alpha myosin heavy chain gene, promoter region.
ACCESSION   AY191158
VERSION   AY191158.1 GI:32493126
KEYWORDS   .
SOURCE   Rattus norvegicus (Norway rat)
  ORGANISM   Rattus norvegicus
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                Rattus.
REFERENCE   1 (bases 1 to 5487)
  AUTHORS   Haddad,F., Bodell,P.W., Qin,A.X., Giger,J.M. and Baldwin,K.M.
  TITLE   Role of Antisense RNA in Coordinating Cardiac Myosin Heavy Chain
                Gene Switching
  JOURNAL   J. Biol. Chem. 278 (39), 37132-37138 (2003)
   PUBMED   12851393
REFERENCE   2 (bases 1 to 5487)
  AUTHORS   Haddad,F., Qin,A.X. and Baldwin,K.M.
  TITLE   Direct Submission
  JOURNAL   Submitted (04-DEC-2002) Physiology and Biophysics, University of
                California Irvine, Medical Science 1, D360, Irvine, CA 92697, USA
FEATURES   Location/Qualifiers
     source   1..5487
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /chromosome="15"
                     /map="15p13"
     mRNA   join(<1..136,703..824)
                     /product="beta myosin heavy chain"
     CDS   join(<1..136,703..720)
                     /codon_start=2
                     /product="beta myosin heavy chain"
                     /protein_id="AAO34581.1"
                     /db_xref="GI:32493127"
                     /translation="EEQANTNLSKFRKVQHELDEAEERADIAESQVNKLRAKSRDIGA
                     KGLNEE"
     misc_feature   825..5024
                     /note="beta-alpha myosin heavy chain intergenic spacer
                     region"
     misc_feature   complement(2868)
                     /note="antisense transcription start site"
     misc_feature   complement(2905)
                     /note="antisense transcription start site"
     gene   5025..>5487
                     /gene="alpha myosin heavy chain"
     TATA_signal   5025..5030
                     /gene="alpha myosin heavy chain"
ORIGIN

Query Match   40.5%; Score 2320.2; DB 10; Length 5487;
Best Local Similarity   77.5%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 583; Indels 425; Gaps 38;

Qy   1 GGATCCTGCAAGGTCACACAAGGGTCTCACCCACCCAGGTGCCCAGTCAGTCTCCAATTCAGT   60
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   576 GGATCCCTCAAGGTCACACAAGGGTCTCACCCACCCAGGTGCCCAGTCAGTCTCCAACTTCAGT   635

Qy   61 TTCCATGCCTTGTTCTCACAATGCTCGGCCCTCCCCCCTCCTCCAGAGCTAATTTGGACTTTGTTTTTAT   120
              ||| |||||| |||||||||||||||||||||||||||||||||| |||||||||||| || |||||||
Db   636 TTCCATGTGCTGTCCCTACAATGTTGGCCCTCCCCAGAGCTAATTTTGAC-TTGGTTTTAT   694

Qy   121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTTGTGCTACCCAGCTCTAAGGTGCCCGTGA   180
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   695 TTCAAAAGGGCCTGAATGAAGAGTAGATCTTGTGTGCACCCAACCCTAAGGGATGCCTGTGA   754

Qy   181 AGCCCTCAGACCTGGACCTTGCAGCCTTTGGAAGCAGACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT   240
              ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   4768 TCAGGCCGGTCAGAGATCTGACAGGGTGGCCTTCCACAGCCTGGGAAGTTCTCAGTGGC   4827
              ||| |||||||||||||||||||||||||||||
Db   149234 T----------------CCTCCCACCGGCCTGGGAAGTTCTCGGTGGC   149203

Qy   4828 AGGAGGTTTCCACAAGAAACACTGGATGCCCCTTCCCTTCCCACTGCCTCCATCTTCC   4887
              ||||||||| ||||||||||||||||||||||| ||||||
Db   149202 AGGATGTTTC------TACTGGATGCCCCTTCCCTTCCCACTGCCTCCCTC----   149156

Qy   4888 TCCTGGGGATGCTCCTCCCCGTCTTGGTTATCTTGGCTTCTTCGCTTCGTCTTCAGCAAGATTTG   4947
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   149155 ---------------CCTTGTCTTGATTAATCTTGGCTCTTAGTGTTCAGAAGAAGATTTG   149112

Qy   4948 CCCTGTGCGTCCACTCCATCTTCTCTACTGTGTCCCGTGCCTTGCCTTCTTGCG   5007
              ||||||||||||||||||||||||||||||||| ||||||||||||||
Db   149111 CCCGGTGCGTGCTACTCCATCTGTCCACTACTCGCTC--------TGCCTTGCCTTCTTGTG   149059

Qy   5008 TGTCCTTCGTTCTTCCACCCATTTCACTTCACCTTTCTCCCCCTTCTCCCCTTCATTTGTATTCAT   5067
              ||| ||||||||||| ||| |||||||||||||||||||||||||||||||||||
Db   149058 TGTTCTCCTTTTCCCACCGTGTTTCCACTCACTGCCTCCCCCCCCCCCTTCATTTTA-   149000

Qy   5068 CCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCTTTCTCCCTTCCTT   5127
              |||||||||||||||||||||||||||||||||||||||
Db   149000 --   148999

Qy   5128 CCTTCCTTCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACC   5187
              |||||||||||||||||||
Db   148999 ----TCCTTCCTTTCTTCGTGTCTGTGTCAGAATGCTGGGAATCAAACC   148959

Qy   5188 TGGGGTTCC--CACCCTATGTAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGCTG   5245
              |||||||| ||||||| ||||| |||||||||||||||||||||||||||||||
Db   148958 CAGGGCTTCATACACGTCAAGTAAGCAATCTCCCAGTGAGCCAAAGCTTTAATCCT-CTG   148900

Qy   5246 GGTGCTCTCTTAACCTTCCTCACCCCCCCGGCTGTCCTGTTCCATCCTGTCAGGATCTCT   5305
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db   148899 GGTGCTGTCTTACCGACCTCCACTCCCGTCGTGTTCCGTGTCCTCTTCCTTCCGTCCAGCTG   148840

Qy   5306 AGATTGGTCTCCAGCCTCTGCTACTCTCTTCCTCTTCTGTCGTCCTCTGTCCAGCTG   5365
              |||||| ||||||||||| |||| ||| |||| |||||||||||||||
Db   148839 -GGTCCGTCTCTCAGCTTCTGCTACTCCTCCAGCTTCTCCTGTCGTCCTCTGTCCAGCTG   148781

Qy   5366 CGCCACGTGGTGCCTCGTTCCAGCTGTGGTCACAATTCTTGGTCCACCATTCTTCAGGATTCTCTGAAAAGTT   5425
              ||||||||||| ||||||||||||||||||||| ||||||||||
Db   148780 CACCTGGTGGGTGCGTGCCTACTTCCACCGGTGTCCAAATTCTCTGAAAAGATTAACCGGG   148721

Qy   5426 AACCAGGTGAGAATGTTCCCTGTAGACAGCAGCAGATTCT-CCCGGAAGGCTCAGGC   5484
              ||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db   148720 TGAGAATGCCCCCCAGTTCCCCTGTAGACAGCAGCAGATCATTCCCGAGCAGCCCAGAC   148661

Qy   5485 TTCCAGCCTCCTCTTTCTTCTGCCCAGCTGCCGCACTCTTAGCAAACCTCAGCCACCCT   5544
              ||||||||||||||||| ||||||||||| ||||||||||||||||||||||
Db   148660 TTCCAGCACCCGC--CCTCTGCCCAGCACCTGCCACTCTTAGCAAACTTCAGCCACCCT   148603

Qy   5545 TACCCCACATAGACC--------------------------   5559
              |||||||||||||||
Db   148602 TCCCCACATAGACACCAAGTCTTGCAGAGAGACTTCTGAGTTTCTGCAA   148543

Qy   5560 -------------TCTGACAGAGAAGCACTTACATGGAGTCCTGGTGGGAGA   5603
                           |||||||||||||||||||||||| |||||||||||||
Db   148542 GGAAAGGAGAACTCTTTGTGATGGGGAGGCAGGAGCAGGCACCTTTACACGGAGTCCTGACGGGAGA   148483

Qy   5604 GCCATAGGCTACGGTGTAAAAGAGGCAGAGGAAGTGTGGTGTAGGAAAGTCAGACTTCA   5663
              ||||||||||||||||||| |||||||||||||||||||||||||||||
Db   148482 GCCATAGGCTATGGCATAGCAGAGAGGGCAGAGGGAGGTGGGTGGAAT-------TGGACTTCA   148432

Qy   5664 CATAGAAGCCTAGCCACCCACACAGACAGAAATGACAGACATCCCTATCCTCCCCCATAGA   5723
              |||||||||||||||||| ||| |||| ||||||||||||||| || |||||||||
Db   148431 CACAGAAG-CTAGCACACACCAGGAAT---GACATATCCTCCTCATCTCCCCCCATATGA   148377

Qy   5724 GTTTGAG   5730
              |||||||
Db   148370 GTTTAAG   148370

```
Db   755  AGCCCTGAGACCTGGAGCTTTGAAACAGCACCTTAGGCAGAAACACAATAAAGCAATTT  814

Qy   241  TCCTTAAAGCCAAAATCCTGCCTCTAGACTCTTCTTCTGACCTCGGTCCCT-GGGCTC  299
          ||||| ||||||||||||||||| ||||||||||||||||||||||| ||||  |||||
Db   815  TCCTTCAAGCCAAAATCCTGTCTCTAGACTCTTCTTCTGACCTCGGTCCCGGTGGGGCTC  874

Qy   300  TAGGGTGGGGAGTGGGGCTTGGAAGAAGGTGGGGAAGTGGCAAAAGCCGATCCCTA  359
          ||||||||||||||||| ||||||||||||||||||||| |||| ||| |||||
Db   875  TAGGGTGGGGAGTGGGGGACTTGAAAGAAGA---GGGGAAGTGGCAAAAGGCCACTCCCCA  931

Qy   360  GGGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGAGATCCTCCTCAGC  419
          || ||||||||||||||| ||||||||| |||| ||||||||||||||||||||| ||||
Db   932  GGACTCTGTGAAGTTCTGAGCCTTCCCAGACACTGGCTGGCTCATAGATCCTCCTTCCAGC  991

Qy   420  CAAACATAGCAAGAAGTGATAACCTCCTTGTGACTTCCCAGGCCCAGTACCTGTCAGGT  479
          |||||| ||||| ||||||||| ||||||||||||||| |||||||||||||| ||
Db   992  CAAACAGCAGGAGAAGTGATGCTTGTTCCTTGTGACTTCCCAGGCCCCAGTGCTGTCCGGT  1051

Qy   480  TGAAACAGAGATTAGAGAGAAGCCCTGAACTCTGAAGCTTCATCACCAAGC---C  539
          ||||||||| |||||||||||||| || |||||| ||||||||||||||||||     |
Db  1052  TGAAACAGGATTTAGAGAGAAGCCTCAAACTCTCAATCTCTCAAAACTCACCAGTGCTTCTAC----C  1107

Qy   540  AAGCACCTAGGTGCCACTGCTAGTGTATCCTCACGCTGATAATAATGCGAGCTG--GGC  597
          |||| | ||||||||| | |||| ||||| |||| || |||| |||  |||||  |||
Db  1108  ACACACCTACATGCCGGTTCTAATCAGTAATCCTACACTGATAATACGCAGAGCTGACCAC  1167

Qy   598  CACAGAAGTCCTGGGGTGTAGAACTGACCAGTGACTTTTCAGTCGGCAAAGGTATGACC  657
          |||||||||||||| |||| |||||| |||| |||||||||| |||| ||||||||
Db  1168  GGGAGAAGGCCCTGGAGTGGAGTGGAGGCGGTCTGCGAGGTATGACCA  1227

Qy   658  CCCTCAGCAGAT--GTAGTAATGTCCCCTTAGATCCCATCCAGGCAGGTCTCTAAGAG  714
          ||||| |||||   ||| |||| | ||||| || ||||| |||||||||||||||||||
Db  1228  CCCTTGGCAGATATGGCAGATGGGGTCCCCTTAGAACCCATCCCAGGCAGTCTCTGAGAG  1287

Qy   715  GACATGGGATGGAGATGAGTCATGTGGCATTCCAAACACACAGCTATCCACAGTGTCCCT  774
          |||| ||||||| |||| |||||| |||| |||||||||| ||||||| |||| |||| |
Db  1288  GACACAGGATGGACAGATGTAGTCATTCCAAACAACACAGCTATCCACAGTGACTTCT  1347

Qy   775  TGCCCCTTCCACTTAGCCAGGAGACAGTAACCTTAGTCTTCTTCCTCCCCCATCC  834
          || ||||||||| |||||| ||| | ||||| ||||| ||||| |||||||| ||
Db  1348  TGTCCCTTCCCATTAGCCATGATAGCAGTAACCTTAGTCTTCATTTCTTCTTCCCCATCT  1407

Qy   835  TCCCAGGACACACCCCTGGTCTTGCAGTATTCATTTCTTCCTTCACGTCCCCTCTGTGAC  894
          |||||||| | |||| |||| | |||| || ||||||||| ||| |||| ||
Db  1408  TCCCAGGACCCAACCCCGTCTATGGTATTCATTTCTT-CTTCATTTCTT-CTTCATTTCCCCACCATGAC  1466

Qy   895  TTCCATTTGCAAGGCTTTTGACCCTCGAGCTGCTCTGGGAAGATAGAGTTGGCCTAGGTG  954
          |||||||||||| ||||||| ||| | | ||||||||| |||||||||||||| |||
Db  1467  TTCCATTTGCAGGGTCTTTTACCTCTGCAGTAGCTGCAGTGAAGATGGGAAGTTTGGCTCTATATG  1526

Qy   955  TGGCAAGCCATCTCAAAGACAATGGCTGGATTGGAAGGCTCTTTGAAGGATCAGGA  1014
          |||| | ||||||| ||||||||| |||| ||||||||||| ||||| |
Db  1527  TGGCTCTGAAAGACATCTTCTTTGAAACATCTTCTTTGAACCTCAGAGATTTTGTTAGG-GGGGAC  1585

Qy  1015  ACTAAATCACTGGCCGCGGGGCTAGAAAAAGAGTTGAGTGAGTCGGCTCAGCTAAG  1074
          |||||||||| |||| ||||||| || |||||||||||||||||| | |
Db  1586  ACTAAATCACAGAGAGGGTCTGGAGTGGTCTGGGGTAGAAAAAAGTGA-------GCTCCAGCTAAG  1637

Qy  1075  CCAAGCTAGTCCCCGAGATACTCTGCGCTGCCCACAGCTGGGCTGCTGCTGGTAGCTTTAGGAATG  1134
          || |||||||| ||||||||||||| |||||| ||| || |||||| ||| |||||||||||||||
Db  1638  CAAAGCCAGAATCCTTTGGGTATT----CACAGCTGGGCTGCT-GAGGTAGGTTTAGGAATG  1692

Qy  1135  TGGGTCTGAAAGACAATGGCTGGATTGGAAGGCTCTTTGAAGTCTCCCCTCAACCCCACCT  1194
          |||| |||||||||||| ||| || ||||| | ||||| ||||||||||| |||||||
Db  1693  TGGCTCTGAAAGAACATCTTTGCCCTGCGTAACCTTGCCCTGTTGTTAGG-GGGGCT  1752

Qy  1195  ACAGACACACTCGTGTGTGGCCAGACTCCTGTTCAACAGCCCCTCGTCTGACCACTG  1254
          || |||||| |||||||| ||||| ||||||||||||||||||||||||||    |
Db  1753  AAATCTCCCCTCAACCCCACCATAGACACGCTTGTCATGGCCAG--------ACTG  1802

Qy  1255  AGCTAGGCAACCAGACAGGCATGATGGCCCTGTTGCTGAGGATGAAGAGTTGGTTACCAATAGCAA  1314
          ||| |||||||| ||||| ||||||||| |||||| ||| || |||||||||||| |||||||||
Db  1803  AGTTAGGCAACTGGACAATGGGCACGATGGGAGGACGAGGGGTTG----CCAATAGCAA  1857


Qy  1315  AAACAGCAGGGGGAGGAGAACAGAGAACAGAGAAATAAGGAAGGAAGAGGAAAGGCCAGTCA  1374
          |||||||||||| ||||||| | |||| || | | ||| ||||| |||| || |||
Db  1858  AAACAGCAGGTGAGGGAGAACAGAGAATGAAATAAGGAAGGACGGAACGGCC----CA  1914

Qy  1375  ATCAGATGCAGTCAGAAGAGATGGAAGCCAACACACAGCTTGAGCAGAGAACAGAGAAA  1434
          ||||||||||||||| |||||||||||| ||| ||||| ||||| ||| | |||| |||
Db  1915  ATCAGATGCAGTCAGAAGAGATGGGAAGCCAATCCAATGCCTGAAGCTTGAACAAAGAAATAGAAAG  1974

Qy  1435  AGGGAGAGA  1443

Db  1975  AGGAAGAGATGGCCAGATCATTAGGGTGGACTTCCTCCATTAATTAAGATGCCTCTGT  2034

Qy  1444  ------TTCTGGGCATAAGGAGGCCACAGAAAGAGAGGCCCAGCCC  1484
                ||||||||||||||||||||
Db  2035  CCGACTCTACATCCCTGGGTACTGGTCATTGCAACTCAGGAATAGCAACAGACAGAGAACT  2094

Qy  1485  CCCAAGTCTCCTCTTTATACCCTCATCCC----  1513

Db  2095  GTATATGGCCATTTCACATTCTCAATCTGAAGCTTGACACACTTGACAATTTCTAGTT  2154

Qy  1514  ----  1513

Db  2155  TACATTGTGTTACTCTATACAGATATGTCTTCATGGACAGTCGTGTGATGCTCCCTGAAT  2214

Qy  1514  ----  1513

Db  2215  ACAAGGTCCTTTTTTATGAAATCAGTGTGCATTAAATCAAGTCTGCATAGGGAGGCCA  2274

Qy  1514  -------GTCTCCCAATTAAGCCCACTCTCTTCTCCTAGATCAGAC  1551
                 ||||||||| ||||| |
Db  2275  CAGGAAGAGCCCAGGTCTCCCCAGGTCTCCCCTTTATGCCCTGGTCCTTAAGCCCATTC  2334

Qy  1552  CTGAGCTGCAGCGAGAGACACCCGTAGGCAGGATCACACTGATGAAGGAGAGTGTTGGAG  1611
          |||| ||   ||||||| ||||| |||
Db  2335  CTCTTC-CTGAACAGACCTGAGCAGTGGAAATCCGTAGAGGTCTAAAGAGGATCA  2393

Qy  1612  AAGTCCAGGGCAACCTAAGAGCCAGCCCAAGAGCCAAGAGCAGATAAAGGTGCTTCAAAG  1671
          |||||||||||||||||||||||| |||||||||| ||||||||||||||||
Db  2394  CAGTGGAAGGAGATGTGTGACCAGACACACAGAGAACAAGAGAGATAAAGGACGCAGCAAAG  2453

Qy  1672  GTGGGCCAGGCTGTG--CACACAGAGGGTCGAGGACTGTTGGAG  1729
          |||| ||||||||   |||||||||||||||||||
Db  2454  GCAGCTAGCCCGTGCACACACAGACACAGAGAGGGTCAAGGACTAGCGGTAGCGGAAGATAAGGA  2513

Qy  1730  TGATGCTCAGAATGGGCGGGGGGGATTCTGGGGGGGGAGAGAAGGTGAGAAGGA  1789
          ||||||||||||||||| || ||
Db  2514  TGATGCTCAGAATGGGTGGTGAGAGAGA  2541

Qy  1790  GCCCTGGAACAGAGAATCTGAAGCGCTGGAAAACGGATCCAATAAAGGGAAGAACCCAGGCT  1849
          |||||||||||||||||| |||| || ||| || ||| |||||||||||||||||||||
Db  2542  ---CGGGACTGAGACCTGGAAGCCCTGAAATACTGTAAAGGGAAAACCCAGGCT  2598

Qy  1850  ACCTTTAGATGTAAATCATGAAAGACAGGAGGAGAGTTGAGTCGGCTCAGCCTGTGAAGGACC  1909
          |||||||||||  ||||| ||||| || |||||||| || |||
Db  2599  ACCTTTAGATGGGAAGTCATGAAAGTCAGGGTCATCAGGGGATCCAGTGAGAGAGTAGGAGAATA  2658

Qy  1910  CCGGGGCAAGCAACATGGAAGCAAGGACAAGCAGGTTGAGCGCTCCGTGAAATCAGCCTGC  1969
          |||||||||
Db  2659  --------AGACAAGGACAAGCAAGCCAGGTTGAGTGCTCTGTGAGATT-------  2695

Qy  1970  TGAAGGCAAGCGCCTGGTATGAGCAGCACCAGAACAGACAGCAGAACCCAGGCTAGGGTCGACA  2029
          || |||||||||||||||| |||||| |||||||||| ||||| |||||| |||||||| ||||
Db  2696  ---CCGGATATGAGCCGGAACAGACAGCAGCGCCCGAACAGACAGCAGAGGCTAGGGCTGATTTAGAGACA  2742

Qy  2030  GGGAACAGAGGTAGACAGGAGACAGACACGGGGACGCCAGTAACAAGGAATG  2089
          |||||| ||||||| |||||||||||||||| |||||||||||| |||||||
Db  2743  GGGAACTGGAAGGTGGGCACAGACAGAGGAACGACAGGGGATCCAGTAACAAGGGATG  2802

Qy  2090  GTCCTTCTCACCCTGTGGCCAGAGCGTCCATCT-GTGTCCACCATACTCTAGAATGTTCATC  2148
          || |||||||||||||||||||||||||||||| |||||| ||| |||| ||||| |||||
Db  2803  GTGCCTTCACCCTGTGGCCAGAGCGTCCACTTCACTTACTCTAGACGTTCATC  2862
```

```
Qy  2149 AGACTGCAGGGCTGGCTGGGAGGCAGCTGGAAAGAGTATGTGAGAGCCAGGGGAGACAA 2208          Db 3927 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 3986
         |||||||||||||||| |||||||        ||||||||||        |||||  ||||||||                   TGGCAGCCTGGAGAGGCCGGAAGAGAGCAGCCCCCACTTCACCCCCTTCTCCATCAGCAC
Db  2863 AGACTGCAGGGCTGGTTGGGAGAGAGTGAGAG-CGGGGGAGACTA 2921                       Qy 3287 CAGGATTAACACCTTCTGGCCTTCCCCTTCCCCACTTCCACCTCCACCTCC 3346
                                                                                           |||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2209 GGGGGCCTAGGAAAGGAAGAAGGAGAGGAGGGAGCAAACCAGGCAAAACCAGGCAGAGCCCAGAA 2268     Db 3987 CAGGATTAACACCTTCTGGCCTTCCCCTTCCCCATCAGGAGTGGAGGGTTGCAG 4046
         |||||||||||||||||||||||||||||||                                        
Db  2922 GGGGACCGGAGGAAAGGAGGGAGAGGGAGCCCACAGAGGGCAGAGCCCAGGA 2981                 Qy 3347 AGGGAGGTAAAAACCTACATGTCCAAACATCATGGTGCACGATATATGGATCAGTATGT 3406
                                                                                           |||||||||||||||||           ||||||||||||||||||||||||||
Qy  2269 CTGAGTTAACTCCTTCCTTGTCGATCTCCATAGGAGGCAGTGGGAACTCTGTGACCAC 2328           Db 4047 AGGGAGGGTAAGAACCTAGAT------AATTGTGCATGATGATATGCACTAGTATGT 4098
         ||||||||||||||     ||||||||||||| ||||||||||||||
Db  2982 CTGATTTAACTCCTTCCTTGT---ATCTTCCATAGGAGGCAATGGAAATCTGTGCCCAC 3038        Qy 3407 GTAGAGAGGCAAGAAAGGAAATCTGCA------GGCTTAACTGGGTTAATGTGTAAA---GT 3457
                                                                                           ||||||||||||||||||||||||||      |||||||||||||||||||||||     ||
Qy  2329 CATCCCCCATGAGCCCCACTACCCAGGTTGGCCTGAGTGGCATTCTAGGTTCC 2388             Db 4099 GTAGAAGCAAGAAGAGAAATCTGAGTGAGTTAATGTGTAAAGTCTCTGTGTGTGCATGTGC 4158
         |||                       |||||||||||||||||||||||||||||||
Db  3039 CAT------CCCCCACTACCTATACCAAGTTGGCCTGAGTAGCATACTAGGTTCC 3088             Qy 3458 CTGTGTGCATGTGTGTGTCTGACTGCAGCGTTGTGTCTGACTGAAAACGGGCATGGCTGTGCAGTTCAGTTCT 3517
                                                                                           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Qy  2389 CTGAGGACAGAGCCTGGCCTTTCTGTCTGGAACCAAGCTGACCCAATGTTCTCA 2448             Db 4159 ATGTGAATGCATGTGTGTGTCTGACTGCAGCGTTGTGTCTGAATGTGCAGTTCAGTTCT 4218
         ||||||||||||||||||||  ||||||||||||||||||||| |||||||||||
Db  3089 CTGAGGACAAAGCCTGGTTTTTGCGCTTGGACCTCAAGTGACCCAATGTTCTCA 3148             Qy 3518 GTCGTGTGAGGTTACCAGCTGCAGGTTTGTGTGTAAATTGCCCAAGGCAAAGTGGTGAA 3577
                                                                                           |||||||||||                        ||||||||||||||||||||
Qy  2449 GTACCTTTGAATGCCCTCAAGGAGCTTGAGAACCAGGCAGCTGACATATTAGGGCCATGGCT 2508     Db 4219 GTATGTGACATTACCAGCTGCAGGTTTGTGTGT------AAGTCAGTGAG 4263
         ||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db  3149 GTGCCTTATCATGCCCTCAAGGAGCTTGAGAACCAGGCAGCTGACATATTAGGCCATGGGCT 3208    Qy 3578 TCCCTTCCATGGTTTAAAGAGATTGGATGATGGCCTGCATCTCAAGGACCATGGAAAAAATA 3637
                                                                                           ||||||||||||||||||||||||| |||||||| |||||||||||||||||||||
Qy  2509 AACCCTGGAGCTTGCACACAGGACCTCAAGTGACCTGCAGACAGG 2568                      Db 4264 ACCCTCCCATGGTTTAAAGAAATTGGATAGCCTGCGTGATGTGCAGCATTGAAAGTA 4323
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  3209 AACCCTGGAGCTTGCACACAGGACCTCAAGTGACCTCCGGGGACACAGACAGG 3268             Qy 3638 GAATGGACACTCTATATGTGTCTCTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGT 3697
                                                                                           || |||||||||||||||||||||| ||||||||||||||||||||||||||
Qy  2569 TGGCCTTTATCCCCAAAGAGCAACCATTTGGCATAGTTGGCTGCAAATGGAATGCAAGG 2628       Db 4324 GAATGTAGACTTTATATGTTCCTAAGCAAAGTAGCAAAGTCTTTGGAGGACACCTGT 4383
         ||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||||
Db  3269 TGGCCTTCATCCTCGGAGCAACCATTTGGCACCAGCAGATGGGAACGCAAGG 3328              Qy 3698 CTAGAGATGTGGGCAACAGAGACTACAGACAGTATCTGTACAGAGTAAGGAGGAGAGGA 3757
                                                                                           ||||||||||||||||||||||||||||| ||||||||||  ||||| |||||||
Qy  2629 TTGAATCAGGTCCCTTCAAGAATACTGCATGCAAGAGACCCTGGAGGAGAGGGGT 2688           Db 4384 CTAGAGATATGGGCAACAGACAATAGACAGTATCTGTACAGAGTAA--AGAGGAGGAGG 4441
         |||||||||||||||||  |||||||||||||  |||||||||||||||||||||||
Db  3329 TTGAAACAGGTCCCTTCAAGAATCAAGATCCGCTACAAGAATCTAAGACCCCTGGAGGAGGA 3388    Qy 3758 GGGGGTGTAGAATTCTTCTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGAT 3817
                                                                                           | |||||||||||||||||||||||||||||||||||||||||||||| |||
Qy  2689 ATGTCCTGCCCCCACCCACCATAAGGGGAGGGGTGAACTATCCTAGGGGGGCTGGCCACTTG 2748    Db 4442 TGAGGTATAGAATTCTCTTACTATCAAAGGGAAACTGAGTCATGCACCTGCAAAATGAAT 4501
         | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3389 ATGTCCTGCCCCCACCCACCCATAGGGGAGGGGTGAACTATCCTAGGGGGCTGCAACCTTG 3448     Qy 3818 GCTCTCCCTAGACATCATGACTTTGTCTCTCGGGGAGCCAGCACTGTGGAACTTCAGGTCT 3877
                                                                                           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  2749 GGGAGACACCACATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCTGTCCTGC 2808        Db 4502 GCCTCCCTGGACATCATGACTTTGTCCCTCGGGGAGCCAGCACTGTGGAACTTCCAGGTCT 4561
         ||||||||||||||||||||||||||||||||||   |||||||||||||||||||||||
Db  3449 GAGAGACACCACATCACCCGAGAGTGCTGACTGACTCAGAAACACTGACCGCCCTAAAGACCCAC 3508  Qy 3878 GAGAGTAGGAGGCTCCCCCTGAGCCTCAGCATGCAGATAGCCAGGTTGAAA--GGG 3935
                                                                                           ||||||||||||||||||||||||||||||||||||||||||||||     |||
Qy  2809 CCACCCTCCACACTCTAGAGCTATATT--GAGAGGTGACAGTAGGGTGGGAGCTGGT 2866         Db 4562 --GAGAGTAGGAGGCACCCCCTGAGCCTGAGCCTGCGTGAGCTAGGGTGTAAAAGAGG 4619
         ||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db  3509 CCACTTCCACACTCTAGAGCTATATTGAGAGAGAGACAGTAGGGTGGGGAGCTGGG 3567          Qy 3936 GGAAGGGAGGGAGGGAGGCCTGGGATGGGAATGCTGCCAAGGAGACAGATATTAAGCCT 3995
                                                                                           |||||||||||||||||||||||||| ||||||||| ||||||||||||||||||||||
Qy  2867 AGCAGGGAGAGTGTTCCTGGGCATGGGCATGGGTGGGGGTGGA 2926                        Db 4620 GGAGGGGGAGGGAGGGTGAATGGGACAGGGGACAGGGGACAAATATTAGGCCCA 4679
         |||||||||||||||||||||||||||||||||||||||||||
Db  3568 AGCAGGGAGAGTGTTCCTGGGTCGGGGGTGGGGGTGGGGGTGGA 3626                       Qy 3996 GGAAGAGAGGTGACCCTACCCAGT-TGTTCAACTCACCCTTCAGATTAAAAATAACTG 4054
                                                                                           |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||
Qy  2927 CTTTGTCTCCTGAACCACAATGTCTACTTAGTTATAACAAAACTGTTAAACAGACCCA 2986        Db 4680 GTAAGAGAAGGTGACCCTTACCCAGTGTGTTCAACTCAGCCTTCAGATTAAAAATAACTA 4739
         |||||||||||||||||||||||||||||||||||||||||||||||
Db  3627 CTTTGTCTCCTGAACCACAATGTCTACTTAGTTATTAACATGCATCATGAAGACCCGGAGACA 3686   Qy 4055 AGGTAAGGGCC---TGGGTAGGGAGGTGGTGTGAGACCTCATG 4111
                                                                                           ||||||||||   ||||||||||||||||||||||||||||||
Qy  2987 ACATCTACGACCTCTGAAAAGACAGCAGCGGGGTTGTCTCTGAGCCTTG 3046                  Db 4740 AGGTAAGGGCCATGTGGGTAGGGAGGTGGTGTGAGACGTCCTGTCTCCTCTATCTG 4799
         ||||||||||||||||||||||| |||||||||||||||||||||
Db  3687 ACACCTACAACCTCTGAAAAAACAGCAGTCCTGGGCGGGGTTGGGGCCTTG 3746                Qy 4112 CCCTTGAGCCCTTTGGGAGGAAGAATGTGCCAAGGAGGACTAAAAAAAAGGCCATGAGCCA 4171
                                                                                           ||                ||||||||||||||||||||||||||| ||||||||||
Qy  3047 GGTGCTTGATGGTGCCACAAAGGAGGGCATGGAGTGTGAGTAAGGCCCAGGAGGCGTTA 3106       Db 4800 CCCATCGGCCCTTTGGGGAGGGAGGAATGTGCCAAGGAGGACTAAAAAAAAGGGCCCTGAGCCA 4859
         ||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db  3747 GGTACTTGATGGTGCCACAAAGGAGGGCATGCATGCATGTGAGTAAGGGGGTTA 3806             Qy 4172 GAGGGGCGAGGGCAACAGACCTTTCATGGCAAACCTTGGGGC 4214
                                                                                           ||||||||||||||||||||||||||||||||||||
Qy  3107 GAGAAGGGCACTTGGGAAGTGGGGTCAGTCTGCGAGCCCCTATCCGAGCCCCTGCTAAGACCCGGAAGTCTGGAGCCT 3166   Db 4860 GAGGGGCGAGGGCAGCAGACCTTTCATGGCAAACCTCAGGGC 4902
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db  3807 GAGAAGGGCACTCGGGAAGGGGGTCAGTCTGCGAGCCCTATCCGAGAGCCCTATCCCAGCAGGAGGGTTA 3866
Qy  3167 GGGGCCAACTGGTTGTAAATCTTCTGGGCCTGCCAGGCATTCAAAGCAGCACCTGCATCCTC 3226
         ||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||
Db  3867 GGGGCCAACTGGTTGTAAATCTTCTGGGCCTGCCAGGCATCCCGAGTAGCACCTGCATCCTC 3926
Qy  3227 TGGCAGCCTGGGGAGGCAGGCGCGAAGGGAGGCAACCCCCCACTTATACCCTTTCTCCCCTCAGCCC 3286
```

RESULT 7
AC115371/c
LOCUS   AC115371   263901 bp   DNA   linear   HTG 10-MAY-2003
DEFINITION Rattus norvegicus clone CH230-118C1, WORKING DRAFT SEQUENCE.
ACCESSION   AC115371

VERSION   AC115371.5 GI:30521556
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE   Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 263901)
AUTHORS   Muzny,D.Marie. Metzker,M.Lee. Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,K., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garza,A., Garza,T., Garza,W.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., McNeil,T.Z., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
            Nwaokelemeh,O., Okwuonu,G., Olarrpunsagoon,A., Pai,S., Pfannkoch,C.,
            Pasternak,S., Paul,H., Perez,A., Perez,A., Pu,L.L.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,M., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,S., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willison,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 263901)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (19-MAR-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 263901)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   On May 10, 2003 this sequence version replaced gi:24954086.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------- Project Information
                Center project name: GHLJ
                Center clone name: CH230-118C1
            -------- Summary Statistics
                Assembly program: Atlas 3.0;
                Consensus quality: 235051 bases at least Q40
                Consensus quality: 237228 bases at least Q30
                Consensus quality: 238719 bases at least Q20
                Estimated insert size: 244896; sum-of-contigs estimation
                Quality coverage: 8x in Q20 bases; sum-of-contigs estimation
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            *   consists of 1 contigs. Gaps between the contigs
            *   are represented as runs of N. The order of the pieces
            *   is believed to be correct as given, however the sizes
            *   of the gaps between them are based on estimates that have
            *   provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *   1 263901: contig of 263901 bp in length.
FEATURES   Location/Qualifiers
     source   1..263901
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-118C1"
     misc_feature   1712..259601
                     /note="clone boundary
                     clone_end:Sp6
                     site:EcoRI
                     end_sequence:BH285254"
     misc_feature   18849..19761
                     /note="wgs_contig"
     misc_feature   complement(155798..262015)
                     /note="clone boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH285252"
     misc_feature   complement(261235..262017)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH285252"
ORIGIN
Query Match   28.9%; Score 1655.4; DB 2; Length 263901;
Best Local Similarity 73.0%; Pred. No. 0;
Matches 2910; Conservative   0; Mismatches 451; Indels 624; Gaps 36;

Qy   1833 AGGGAGAACCAGGTACCTTTAGATGTAAATCATGAAAGACAGGAG-AAGGGAAGCT 1891
           |||| | ||||| |||||||||| | ||||||||| ||||||||| |||||||||
Db 263901 AGGGAATACCCAAGCTACCTTTAGATGGAAGTCATGAAAGTCATGAAAGTCAGGGAGAAAGGGAAGCT 263842

Qy   1892 GGGAGAGTAGAAGGACCCCGGGGCAAGCAAGCATGGAAGCAAGCACAAGGACAAGCCAGGTTGAGCGC 1951

```
Db  262813  TGACCCGCTAAAGACCCAACACCTACAACCTCTGAAAAACAGCAGTCCTGGAGGACGGG  262754

Qy    3029  GGTTGTCTCTGAGCCTTGGGCCTTGATGTGTCGCCACAAAGGAGGGCCATGAGTGTGAGTAT  3088
            ||| || |||||| |||||||||||| || ||||| ||||| ||||||||||||||||
Db  262753  GGTTGCCGTTGGGCCTTGGGTACTTGATGTGTCGCCACAAAGGAGGGCCATGAGTGTGAGTAT  262694

Qy    3089  AAGGCCCCAGGAGCGTTAGAGAAGGCACTTGGGAAGGGGTCAGTCTGCAGAGCCCCTAT  3148
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262693  AAGGCCCCAGGAGCGTTAGAGAAGGCACTTGGGAAGGGGTCAGTCTGCAGAGCCCCTAT  262634

Qy    3149  CCATGGAATCTGAGCCTTGGGCCAACTGTGTGTAAATCTTCTGGGCCTGCCAGGCATTCAA  3208
            ||||||||||||||||||||||| |||||  ||||||||||||||||||||||||||
Db  262633  CCATGGAATCTGGACACAGGGGCCAACTGTGTGTAAATCTCCTGGGCCTGCCAGGCATCCCG  262574

Qy    3209  AGCAGCACCTGCATCCTCCTGGCAGGCTGGCTGGGAGGCAGCTGGAAGGAGCTGAAGAGTATGT  3268
            |||||||||||||| ||||||||| ||||||||||||||||||||| |||||||||
Db  262573  AGTAGCACCTGCATCCTCTGGCAGCCTGGCTGGGAGGCGGCTGGAAGAGACCCCCCACTTACA  262514

Qy    3269  CCCTTTCTCCCTCAGCCCCAGGATTAACACCTCTCGGCCTTCCCCTTCCCCATCCCCATC  3328
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  262513  CCCCTTCTCCATCAGCACCAGGATTAACACCTCTCGGCCTTCCCCTTCCCCATCCCCATC  262454

Qy    3329  AGGAGTGAGGGGTTGCAGAGGGGAGGTAAAAACCTACATGTCCAAACATCAGTGTGCACG  3388
            |||||||||||||||||||||||||||||||||||||          |||||||
Db  262453  AGGAGTGAGGGGTTGCAGAGGGGAGGTAAGAACCTAGAT------AATTGTGTCATG  262402

Qy    3389  ATATATGGATCAGTATGTGTAGAGCGCAAGAAAGAAAATCTGCA------GGCTTAACTGG  3442
            |||||||||||||||||||||||||||||||||||||||||||      |||||||||||
Db  262401  ATATATGACTAGTATGTGTAGAGCAAGAAAGAAAATCTGGAGTTAATGTGTAAAGTCT  262342

Qy    3443  GTTAATGTGTAAA---GTCTGTGTGCATGTGTCTGACTGAAAACGGGCATGGCT  3499
            |||||||||||||   |||||||||||||||| ||| ||| ||||||||||||
Db  262341  CTGTGTGTCATGTGCATGTGAATGCATGTGTGTCTGACTGGAT  262282

Qy    3500  GTGCAGCTGTTCAGTTCTGTCGCGTGATGTTACCAGACTGCAGGTTTGTGTGTAAATTGCC  3559
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262281  GTGCAGCTGTTCAGTTCTGTATGTGACATTACCAGACTGCAGGTTTGTGT-----  262230

Qy    3560  CAAGGCAAAGTGGGTGAATCCTTCCATGGTTAAAGAGATTGGATGAATGGCCTGCCATCT  3619
            ||||||||||||||||||||||||||||||||||||||||||||||
Db  262229  ------AAGTCAGTGAGACCCTCCCATGGTTAAAGAAATTGGATGAATAGCCTGCATGT  262177

Qy    3620  CAAGGACCATGGGAGAAATAGAATGGACACTCTATATGTCTAAGCTAAGGCATGGCAAGG  3679
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  262176  CAGGGAACATTGAAAGTAGAATGTAGACTTTATATATGTTCCTAAGCCAAAGTAGCAAGG  262117

Qy    3680  TCTTTGGAGGACACCTGTCTAGACATGTGGGCAACAGAGACTACAGTATCTGTACA  3739
            |||||||||||||||||||||||||||| ||||||||||||||||||||||
Db  262116  TCTTTGGAGGACACCTGTCTAGAGATGTGGCAACAGAGACTATCTGTACA  262057

Qy    3740  GAGTAAGGAGAGAGGGGTGTAGAATTCTTACTATCAAAGGGAAACTGAGTCG  3799
            ||||||     |||||||||||||||||||||||||||||||||||||||
Db  262056  GAGTAA-----AGAGAGGAGGTGAGGTATAGAATTCTTACTATCAAAGGGAAACTGAGTCA  261999

Qy    3800  TGCACCTGCAAAGTGGATGCTCCTCCCTAGAGTGCTGAGCCCCAGCAGCA  3859
            |||||||||| |||| |||||||||||||||||||||||||||||||
Db  261998  TGCACCTGCAAAATGAATGCCCTGGGGGAGTGCTGAGCCCCAGCAGCA  261939

Qy    3860  CTGTGGAACTTCAGGTCTGAGAAGGTCTGAAGCTGAAGCCTGTATGCAGATA  3919
            |||||||||||||| || ||||||||||||||||||||||||||
Db  261938  CTGTGGAACTCCAGGTCT--GAGACTAGGAGGCACCCCCAGCCTGTGCAGATA  261881

Qy    3920  GCCAGGGTTGAAA----GGGGAAGGGAGGGGAGCTCGGAGTGGAATGGGAGGCAG  3977
            ||||||||||||    ||||||||||||||||||||||||||||||||
Db  261880  GCTAGGGTGTAAAGAGGGGAGGGAGGGGAGCTTGTGTTGGGAGACAG  261821

Qy    3978  GGGACAGATATTAAGCCTGAAGAGAAGGTGACCCTTACCCAGT-TGTTCAACTCACCCT  4036
            |||||||| ||||||||||||||||||||||| ||||||||  |||||||||||
Db  261820  GGGACAAATATTAGGCCCGTAAGAGGTAAGGTGACCCTTACCCAGTGTGTTCAACTCAGCCT  261761

Qy    4037  TCAGATTAAAAATAACTAACTAAGGTAACGGC----TGGGTAGGGGAGGGGAGGGGAGGGCTC  4093
            |||||||  |||||| ||||||||| || |    |||||||||||||||||||||
Db  261760  TCAGATTAAAAATAACTAACGGTAACGGCCATGTGGGTAGGGGAGGGGAGGGGAGACGGGTC  261701
```

```
Db  263841  ------AGACAAGGACAAGCCAGGTTGAGTGC  263798
            |||||||| |
Qy    1952  TCCGTGAAATCAGCCTGCTGAAGGCAGCAGCCCTGGTATGAGACAGCAGCAGAGGCT  2011
            ||||||| | ||||
Db  263797  TCTGTGAGATT--------CCGATATGAGCCGCGGAACAGCAGCAGAGGCT  263758

Qy    2012  AGGGTTAATGTCAGAACAGGGAACAGGGTAGAACAGACAGAGACGGGGGAG  2071
            |||||||||||||||||||||||||||||||||||||||||||||
Db  263757  AGGGCTGATTTAGAGACAGGGAACTGAACGTGGCACAGGAGGTGGCACAGAGACAGGGGAT  263698

Qy    2072  CCAGGTAACAAAGGAATGTCCTTCTCCACCTGTGGCCAGAGCGTCCATCT-GTGTCCACA  2130
            ||||||||||||| ||||||||||||||||| ||||||||||||||  |||||||
Db  263697  CCAGGTAACAAACGGGATGGTGCCTCTCCACCTGTGGCCAGAGCGTCCATCTGGGGGTCACT  263638

Qy    2131  TACTCTAGAAATGTTCATCAGACTGCAGGGCTGCAGGGCTTGGAGGAGAGTATGT  2190
            ||||||||||||||||||||||||||||| ||||||||||||||||||||
Db  263637  TACTCTAGAAACGTTCATCAGACTGCAGGGCTGTGTTTGGGAGGGGCTGAAGAGTATGT  263578

Qy    2191  GAGAGCCAAGGGAGACAGAGGGGCCTAGGAAAGAGAGGGCAAACCAGGCCACACA  2250
            |||||||||||||||||||||| ||||||||||||||||||||||||||
Db  263577  GAGAG-CGGGGGACACTAGGGGCACCGAGGCAAAGGCGAGAGGGCAAACCAGGCCACACA  263519

Qy    2251  AGAGGGGCAGAGCCCAGAACTGAGTAACTCCTCCTTGTCGCATCTTCCATAGGAGGCAG  2310
            |||||||||||||||||||||||||||||||||  |||  |||||||||||||
Db  263518  AGAGGGGCAGAGCCCAGAACTGATTTAACTCCTTGT---ATCTTCCATAGGAGGCAA  263462

Qy    2311  TGGGAACTTCGTGACCACCATCCCCCATGCAGCCCCACTACCCAAGTTTGGCCTG  2370
            ||||||||||||||||||||  |||||||||||||||||||||||||||
Db  263461  TGGAAACTCGTGTGCCACCAT-------CCCCCACTATACCAAGTTTGGCCTG  263412

Qy    2371  AGTTGGCATTCTAGGTTCCCTGAGGACAGAGCCTTGTCTTCTGTCGTCTGACCCAA  2430
            |||||||||||||||||||||||||||||| |||||||||||||||||||
Db  263411  AGTAGCATACTAGGTTCCCTGAGGACAAAGCCTGTTTTTGTCGTCTGACCCAG  263352

Qy    2431  GCTGACCCAAATGTTCTCAGTACCTTGAATGCCCTCAAGAGCTTGAGAACCCAGCAGTGA  2490
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263351  GCTGACCCAAATGTTCTCAGTCCTTATCATGCCCTCAAGAGCTTGAGAACCCAGCAGTGA  263292

Qy    2491  CATATTAGGCCATGGCGTCTAACCCTGGAGCTTGCACACAGACCTCAGG  2550
            |||||||||||||||  ||||||||||||||||| ||||||||||||
Db  263291  CATATTAGGCCATGGCCTAACCCTTGCAGCTTGCACACAGACCTCAAGTTTGGCCTCCGGG  263232

Qy    2551  GACACAGCTGCAGACAGGTGGCCTTTATCCCCAAAGAGCAACCATTTGGCCATAGGTGGCT  2610
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263231  GACACAGCTGCAGACAGGTGGCCCTTCATCCTCGGAAAGAGCAACCATTTGCCAGAGGCACCA  263172

Qy    2611  GCAAATGGGAAGTTGAATCAGGTCCCTTCAAGAATCAGTCCATGCACGACCTAAG  2670
            || |||| ||||||||||||||||||||||||||||||  |||||||||||||
Db  263171  GCGAATGGGAACCGCAAGGTTGAAACAGGTCCCTTCAAGAATACCGGTACAAGATCTAAG  263112

Qy    2671  ACCCCTGGAGAGAGGGGGTATGCTCCTGCCCCCCAAAGAGCCACCATTGGAGTGACTATCCT  2730
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  263111  ACCCCTGGAGAGAGACTGTATGCTCCTGCCCCCACCCCACCCATAAGGGAAGTGAAGTATCCT  263052

Qy    2731  AGGGGGGCTGGCGACCTTGGGGAGACACCACCACCACCACATTACTGAGCCCCAGAAAAACT  2790
            ||||||| |||| || ||||||||||||||||||| ||||||||||||||
Db  263051  AGGGGGGAGTGGGGCTGCAACCTTGGAGAGACACCACCACCACCATCCGCGAGTGCTGTACTCAGAAACACT  262992

Qy    2791  GACCCCCTGTCGTGTGCCCACCACTCCACACTCCACAGCTCAGAGGTCACAGTA  2848
            |||||| ||||||||||||||||| |||||||||||| ||||||||||
Db  262991  GACCCGCTGTGTGTGCCCACCACTCCACACTCTAGAGCTCAGAGAGGTGACAGTA  262932

Qy    2849  GATAGGGGTGGGAGCTGGTAGCAGGGAAGTGTTCCTGGGGTGTGAGGGGTGTAGGGGAAAGC  2908
            |||| ||||||||||||||||||||| |||||||||||||||||||||||||
Db  262931  GAT-GGGGAGGGGAGCTGGGAGCAGGGAAGTGTTCCTGGGGTGTGGGGTGTGA-GGGAAAGT  262874

Qy    2909  CAGAGCCAGGGAGTCTGGCTTTTGTCTCCTGAAACACAATGTCTACTTAGTTATAACAGGCA  2968
            ||||||||||||||||||||| ||||||||||||||||||||||||| |||||
Db  262873  CAGAGCAGGGAGTCTGGCTTGGTCTGGCTTTTGTCTCCTGAACACAGTGTCTACTTAGTTATAACATGCA  262814

Qy    2969  TGACCTGCTAAAGACCCAACATCTACGACCTCTGAAAAAGACAGCAGCCCTGGAGGACAGG  3028
            |||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db                                                                    3028
```

The following is a nucleotide sequence alignment (Qy = query, Db = database) presented in rotated format on the page.

```
Qy   4094  CTGTCTCCTCTGCATGCCCTGAGGCCCTTTGGGGAGGAGGAATGTGCCCAAGGACTAA                4153
           |||||||||||||||||||||||  ||||||||||||||||||||| |||||||||||||
Db 261700  CTGTCTCCTCTCTATCTGCCCATCGGCCCCTTTGGGGAGGAGGAATGTGCCCAAGGACTAA              261641

Qy   4154  AAAAAGGGCCATGAGCCAGAGGGCGAGGGCAACAGACCTTTCATGGGCAAACCTTGGGG                 4213
           |||||||||||  |||||||||||||||||||||||||||||||||||||||||
Db 261640  AAAAAGGGCCCAGAGGGGCAGAGGGCGAGGGCAGCAGACCTTTCATGGGCAGACCTCAGGG              261581

Qy   4214  CCCGTAGTGATCGATTGACAAGAACTCGGCCAATCGATACCCTTCTTCTTCTCTAACGGACAG             4273
                                   |||||| | ||||||||||||||||||| ||||||
Db 261580  C-----------------------TGCTGTCCTCTGCTAACCTCACTCAGGACGCC                  261553

Qy   4274  GAGGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTGTTTAC                   4333
           |||| |||||||||||||||||||||||||||||||||||||||
Db 261552  AAGGGATCAAAGAGGAGGAGGAGCACAGGAGGACAGGAGGGATGGGAGGGAG----------             261507

Qy   4334  CACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATA                      4393
           |
Db 261506  ------------------------------------------------------------             261507

Qy   4394  GAGAAAGTGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGA          4453
           |
Db 261506  ------------------------------------------------------------             261507

Qy   4454  GTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCCCTATCA                      4513
           |
Db 261506  ------------------------------------------------------------             261507

Qy   4514  GTGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAA                      4573
           |                                                         |||||
Db 261506  -----------------------------------------------------TCTGACAGAGAGCAGGCAGGCACTTTACA  261457

Qy   4574  AGTCGAGCTCGGTACCAGCAGAGACTCCAAATTTAGGCAGCAGGCATCATATGGGGAT                  4633
           |||                             |||||||||||||||
Db 261506  -----------GGTCCCAGCAGATGACTCCAAATTTAGGCAGCAGGCACGTGGAATGAGCT              261457

Qy   4634  ATAAAGGGGCTGGAGCACTGAGAGCTGTCAG----AGATTTCTCCAACCCAGGTAAGAGG               4689
           |||||||||||||||||| |||||||||||      |||||| || |||||||||| |||
Db 261456  ATAAAGGGGCTGGAGCGCTGAGAGCGCTGTCGAGAGTGGACTTCACACAGAAG-CTAGCACACACCACCAGGAAT  261397

Qy   4690  GAGTTTCGGGTGGGGGCGCTCTTCACCACCAGGGGGGCTCTCCCACCTAGAGGGAAACTGC              4749
           |||||||||||||||| ||||||||||||||||||||
Db 261396  GAGTTTAGCGTGGGGGCGCTCTCCCACCCGCGCCTCTCCCACCTAGAGGGGAAAGTGT                 261337

Qy   4750  CTTTCCTGGAAGTGGGGTTCAGGCGCGGTCAGGAGATCTGACAGGGTGGCCTTCACCAGCC              4809
           |||||||||||||||||||| |
Db 261336  CTTTCCTGGAAGTGGGGGCT------------------------------CCTTCCCACCGGCC           261305

Qy   4810  TGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAGAAACACTGAGCCCCTTCCCTTACG               4869
           |||||||||||||||||| ||||||
Db 261304  TGGGAAGTTCTCGGTGGCAGGATGTTTCTA---------------CTGGATGCCCCTTCCCTTCCA         261254

Qy   4870  CTGTCTTCTCCATCTTCCCTCGGGGATGCCTCTCCCCCGTCTTGGTTTATCTTGGCTCTT               4929
           |||||| |||||| |                             ||||||| |||||||||
Db 261253  CTGCCTCCTCCTC-----------------------------------CCTTGTCTTGATTAATCTTGGCTCTT  261214

Qy   4930  CGTCTTCAGCAAGATTTGCCCGTGCTGTCCGGTGCTGTCTCTACTGTCTCTACTCTC------            4989
           ||| || |||||||||||| ||||||||||||||| |||||||||||||||||
Db 261213  AGTGTTCAGAAAGATTTGCCCGGTGCTGTCTACTTCCATCTGTCTCTACTCTC                     261160

Qy   4990  TTGGCTTGGCTTCTTGCTTCTTCCTGTGTCCTTCCTTCCTCCATTTCTCACCCTTTTCTCCC            5049
           |||||||||||||||| |||||||||||||||||||||||| |||||| |||||||||||||||
Db 261159  -TGCCTTGGCTTCTTCGTGTGTGTTCTCCTTTTCCACGTGTTTCCACTCCACTGCCTCCCCCC           261101

Qy   5050  CTTCTCATTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCC                        5109
           |||||||||||||| ||||| ||||||||||||||||||||||
Db 261100  CCCCCCCTTCATTTTTATCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCC---------             261070

Qy   5110  TTCCTTTCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTGTGTCAG             5169
           |||||||                                              |||||||||
Db 261069  ------------------------------------------------------CTGTGTCAG           261061
```

```
Qy   5170  AGTGCTGAGAATCACACCTGGGGTTCCCACCC--TTATGTAAACAATCTTCCAGTGAGCC               5227
           ||||||||||||||||||| |||||| |||||    || ||||||||||||||||||||||
Db 261060  AATGCTGGGAATCAAACCCAGGGGCTTCATACACGTCAAGTAAGCAATCTCCCAGTGAGCC              261001

Qy   5228  ACAGCTCAGTGCGTGCGGGTGCTCCTCTTACCTTCCTCCTCTACCTCGTGTCCTGTTCC                5287
           ||||||| |     ||||| |||||||||||   ||||||||||| |||||||| ||||||
Db 261000  AAAGCTTTAATCCT-CTGGGGCTGTCTTACCGACCTCCACTCACTCCCGTGTCTTGTGTCC             260942

Qy   5288  ATCCTGGTCAGGATCTTAGATTGGTCTCCAGCCTCTGCTACTCCTCTTCCTGCCTGTT                 5347
           |||||| |||| ||||| ||||| |||||||||||||| |||||| | |||||||||||
Db 260941  GTCCTAGTCAGGATCTCT-GTCCGGTCTCTCAGCTTCTGCTACTCCTCCTCCTGCCTGCT               260883

Qy   5348  CCTCTCTGTCCAGCTGCCGCCACTGTGGTGCCTCGTTCAGCTGTGGTCCACATTCTTC                 5407
           |||||| |||||||||| | ||||||||| |||||||||||||| |||||||
Db 260882  CTTCTCTCCGTCCAGCTGCACTCCTGTGGTGCCTCCTCCTTCCAGCTCCCAAATTCTCT               260823

Qy   5408  AGGATTCTGTCGAAAAGTTAACCAGGTGAGAATGTTTCCCCTTGGGTGAGATCACAGAT                5467
           || ||| ||| |||||||||||||| ||||||||||||||| |||||||| | ||||||
Db 260822  GTGAAAAGATTAACCGGGTGAGAATGCCCCCCAGTTTCCCCTGTAGACAGCAGCAGACATCATGAT        260763

Qy   5468  TCT-CCCGGAAGTCAGGCTTCCAGCCTCTCTTTCTCTGCCCAGCTGCCCCGGCACTCTTA               5526
           ||| |   ||||||||||||||||| ||||  |||  | |||||||| |   |||||||||
Db 260762  TTTCCCCAGAGAAGCCAGACTTCCAGCCCGC--CCTCTGCGCCCAGCAACCTGGCACTCTTA            260705

Qy   5527  GCAAACCTCAGGCACCCTTACCCCACATAGACC                                          5559
           |||||||||||| |||| |||||||||||
Db         ------------------------------------------------------------

Qy   5560  ----------TCTGACAGAGAGCAGGCAGGCACTTTACA                                    5585
                      |||||
Db 260704  GCAAACTTCAGCCACCTTCCCCACCATAGACCAAGTCTTGCGAGAGAGCCTCCTTCAGAT               260645

Qy   5586  TGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGAGGCAGGGAAGTGGTCGTGTG             5645
           ||||||||| |||||||||||||||| ||||| ||||||| ||||||||||||||||||||
Db 260584  CGGAGTCCTGACGGGAGAGCCATAGGCCATAGGCACCAGAGGCAGGGAGGTGGTCGGAAT              260525

Qy   5646  AGGAAAGTCAGGACTTCACATAGAAGCCTTCACATAGAAGCCCACCACCACCAGATCCCT               5705
                                  |||||||| ||| ||||||||||||||||  |||| ||||
Db 260524  --------------------TGGACTTCACACAGAAG-CTAGCACACACCACCAGGAAT---GACATATCCCT 260479

Qy   5706  CCTATCTCCCCATAAGAGTTTGAG                                                   5730
           |||||||||||||||||||||||
Db 260478  CCTATCTCCCCATAAGAGTTTAAG                                                   260454
```

RESULT 8

HAMSHCA

```
LOCUS       HAMSHCA                 32415 bp    DNA     linear   ROD 17-FEB-1995
DEFINITION  Mesocricetus auratus alpha-cardiac myosin heavy chain gene exons
            1-39, complete cds.
ACCESSION   L15351
VERSION     L15351.1  GI:402373
KEYWORDS    .
SOURCE      alpha-cardiac myosin heavy chain.
            Mesocricetus auratus (golden hamster)
ORGANISM    Mesocricetus auratus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Mesocricetus.
REFERENCE   1  (bases 1 to 32415)
AUTHORS     Wang,R., Sole,M.J., Cukerman,E. and Liew,C.C.
TITLE       Characterization and nucleotide sequence of the cardiac
            alpha-myosin heavy chain gene from Syrian hamster
JOURNAL     J. Mol. Cell. Cardiol. 26 (9), 1155-1165 (1994)
MEDLINE     95115033
PUBMED      7815459
COMMENT     Original source text: Mesocricetus auratus (strain F1B) male adult
            liver DNA.
FEATURES             Location/Qualifiers
     source          1..32415
                     /organism="Mesocricetus auratus"
                     /mol_type="genomic DNA"
                     /strain="F1B"
                     /db_xref="taxon:10036"
```

```
NQKKKMEADLTQLQTEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERM
KKNMEQTIKDLQHRLDEAEQIALKGGKKQLQKLEARVRELENELEAEQKRNAESVKGM
RKSERRIKELTYQTEEDKKNLVRLQDLVDKLQLKVKAYKRQAEEAEEQANTNLSKFRK
VQHELDEAEERADIAESQVNKLRAKSRDIGAKQKMHDEE"
```

Feature	Coordinates	Qualifier
intron	5364..6525	/number=3
exon	6526..6669	/number=4
intron	6670..6892	/number=4
exon	6893..7049	/number=5
intron	7050..7167	/number=5
exon	7168..7195	/number=6
intron	7196..7411	/number=6
exon	7412..7523	/number=7
intron	7524..7760	/number=7
exon	7761..7853	/number=8
intron	7854..8539	/number=8
exon	8540..8603	/number=9
intron	8604..8869	/number=9
exon	8870..8968	/number=10
intron	8969..9336	/number=10
exon	9337..9440	/number=11
intron	9441..9530	/number=11
exon	9531..9669	/number=12
intron	9670..10083	/number=12
exon	10084..10352	/number=13
intron	10353..11048	/number=13
exon	11049..11219	/number=14
intron	11220..12538	/number=14
exon	12539..12848	/number=15
intron	12849..13362	/number=15
exon	13363..13433	/number=16
intron	13434..13888	/number=16
exon	13889..13976	/number=17
intron	13977..14059	/number=17
exon	14060..14177	/number=18
intron	14178..14298	/number=18
exon	14299..14422	/number=19
intron	14423..14588	/number=19
exon	14589..14725	/number=20
intron	14726..17337	

```
                  /sex="male"
                  /tissue_type="liver"
                  /dev_stage="adult"
misc_difference   1..60
                  /standard_name="overlapping region"
                  /note=", "in misc_difference; The first 60 bp of this
                  sequence overlap with the last 60 bp of the sequence under
                  Genbank accession number L12104. Their combined sequence
                  spans a 66,318 bp genomic region."
                  /replace=""
CAAT_signal       4000..4005
TATA_signal       4032..4037
5'UTR             join(4064..4083,4519..4689,5146..5363)
                  /standard_name="5' untranslated region"
                  /note="The 5' UTR for hamster cardiac alpha-myosin heavy
                  chain gene contains the first two exons and the first 17
                  bp of the third exon; putative"
                  /function="5' UTR for hamster cardiac alpha-myosin heavy
                  chain gene"
exon              4064..4083
                  /number=1
intron            4084..4518
                  /number=1
exon              4519..4689
                  /number=2
intron            4690..5145
                  /number=2
exon              5146..5363
                  /number=3
CDS               join(5163..5363,6526..6669,6893..7049,7168..7195,
                  7412..7523,7761..7853,8540..8603,8870..8968,9337..9440,
                  9531..9669,10084..10352,11049..11219,12539..12848,
                  13363..13433,13889..13976,14060..14177,14299..14422,
                  14589..14725,17338..17480,18082..18255,18431,
                  18950..19085,19295..19385,21101,21490,21820..21946,
                  22037..22155,22444..22640,23183..23366,23503..23668,
                  23762..23886,24640..24948,25332..25545,27609..27734,
                  27843..28118,29407..29502,29617..29751,30113..30136)
                  /codon_start=1
                  /note="putative"
                  /product="alpha-cardiac myosin heavy chain gene"
                  /protein_id="AAB59701.1"
                  /db_xref="GI:402374"
                  /translation="MTDSQMADFGAAAEYLRKSEKERLEAQTRPFDIRTECFVPDDKE
EFVKAKIVSREGGKVTAETENGKTVTVKEDQVMQQNPPKFDKIEDMAMLTFLHEPAVL
YNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVAAYRGKKRSEAPAHIFSISDN
AYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFASIAAIGDRSKKDNPNANKGTLE
DQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIF
QLKAERNYHIFYQILSNKKPELLDMLLVTNNPYDYAFVSQGEVSVASIDDSEELLATD
SAFDVLGFTAEEKAGVYKLTGAIMHYGNMKFKQKQREEQAEPDGTEDADKSAYLMGLN
SADLLKGLCHPRVKVGNEYVTKGQSVQQVYYSIGALGKSVYEKMFNWMVTRINATLET
KQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFFNHHMFVLEQEEYKKEGIE
WEFIDFGMDLQACIDLIEKPMGIMSILEEECMFPKATDMTFKAKLYDNHLGKSNNFQK
PRNVKGKQEAHFSLVHYAGTVDYNIIGWLEKNKDPLNETVVGLYQKSSLKLMATLFST
YASADAGDSGKGKGKKGGSSFQTVSALHRENLNKLMTNLRTHPHFVRCIIPNERKA
PGVMDNPLVMHQLRCNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDS
RKGAEKLLSSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRIITRIQAQARGQL
MRIEFKKMVERRDALLVIQWNIRAFMGVKNWPWMKLYFKIKPLLKSAETEKEMANMKE
EFGRVKESLEKSEARRKELEEKMVSLLQEKNDLQFQVAEQDNLNDAEERCDQLIKNK
IQLEAKVKEMTERLEDEEEMNAELTSKKRKLEDECSELKKDIDDLELTLAKVEKEKHA
TENKVKNLTEEMAGLDEIIAKLTKEKKALQEAHQQALDDLQAEEDKVNTLTKSKVKLE
QQVDDLEGSLEQEKKVRMDLERAKRKLEGDLNVTQESIMDLENDKLQLEEKLKKKEFD
ISQQNSKIEDEQALALQLQKKLKENQARIEELEEELEAERTARAKVEKLRSDLTRELE
EISERLEEAGGATSVQIEMNKKREAEFQKMRDLEEATLQHEATAAALRKKHADSVAE
LGEQIDNLQRVKQKLEKEKSEFKLELDDVTSNMEQIIKAKANLEKVSRTLEDQANEYR
VKLEESQRSLNDFTTQRAKLQTENGELARQLEEKEALISQLTRGKLSYTQQMEDLKRQ
LEEEGKAKNALAHALQSARHDCDLLREQYEEEMEAKAELQRVLSKANSEVAQWRTKYE
TDAIQRTEELEEAKKKLAQRLQDAERAVEAVNAKCSSLEKTKGRLQNEIEDLMVDVER
SNAAAALDKKQRNFDKILAEWKQKYEESQSELESSQKEARSLSTELFKLKNAYEESL
EHLETFKRENKNLQEEISDLTEQLGEGGKNVHELEKVRKQLEVEKMELQSALEEAEAS
LEHEEGKILRAQLEFNQIKAEIERKLAEKDEEMEQAKRNHLRVVDSLQTSLDAETRSR
NEALRVKKKMEGDLNEMEIQLSQANRIASEAQKHLKNAQAHLKDTQLQLDDALHANDD
LKENIAIVERRNTLLQAELEELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLI
```

```
exon        /number=20
            17338..17593
intron      /number=21
            17594..17839
exon        /number=21
            17840..18082
intron      /number=22
            18083..18254
exon        /number=22
            18255..18431
intron      /number=23
            18432..18949
exon        /number=23
            18950..19095
intron      /number=24
            19096..19294
exon        /number=24
            19295..19385
intron      /number=25
            19386..21100
exon        /number=25
            21101..21490
intron      /number=26
            21491..21819
exon        /number=26
            21820..21946
intron      /number=27
            21947..22036
```

```
Query Match         25.2%;   Score 1444.8;   DB 10;   Length 32415;
Best Local Similarity   69.3%;   Pred. No. 0;
Matches   2854;   Conservative   0;   Mismatches   932;   Indels   330;   Gaps   52;
```

```
Qy    226 AGAATAAAGCAATTTTCCTTAAAGCCAAAATCCTGCCTCCAGACTCTTCTTCTCTGACCT 285
          ||||||||||||||||||||||||||| ||||||  |||||||| |||| || ||||| |
Db      1 AGAATAAAGCAATTTTCCTTGAAGCCGAGACCCTGCCTCCAGTCTCTTCTCTTCGCTGTCTT 60

Qy    286 CGGTCCCTGGCTCTAGGGTGGGGAGGTGGGGCCTGGAAGAAGAAGAGGTGGGAAGTGGCA 345
          |||||||||||||  |||||||||  ||||||||||||| ||||||||||| |
Db     61 CCATCCTGGGGCTTTGAGTGGGGAGGTGGGGCCTGAAAGAAGAAGGGGTGTCTTCAAGAG 120

Qy    346 AAAGCCGATCCCTAGGGCCCTGTGAAGTTCGGAGCCTTCCGTGTACAGCACTGGCTCATA 405
          ||     ||||| | |||| |||||| |||||||||||| |||||||| || || |
Db    121 CAAAGGCTCCCCCCAGTCCCTGTGAGCTCTGAGCCTCCTGGCACAGCATTGGTGCACG 180

Qy    406 GATCCTCCTCCAGCCAAACATAGCAAGAAGTGATACCTCCTTTGTGACTTCCCCAGGCCC 465
          |||||||||||||||||| |||| |||||||||  ||||||||| |||
Db    181 GACCCCCCTCCA-CAGAATCAGCAAGAGGTGGCACACCGCT-----CCTTCCCCGGCC 234

Qy    466 AGTACCTGTCAGGTTGAAAACAGATTTAGAGAAGCCTCGAACTCACCTGAACTCTGAAG 525
          |||||||| ||||||||  |||||||||||||||||| |||||||||||| |||||| |
Db    235 AGGGGCCTGTCAGGATGAAA-TGGGTTCAGAGGGGGCTCGAGAGGCCTGTGAACTCATCTGGCGTGCTGAG 293

Qy    526 CTCATCCACCAAGCAAGCACCTAGCTGCCACTGCTAGTTAGTATCCTACGCTGATAATAT 585
          |||||||||  |||||||| ||| ||  ||||||| ||||||| | ||| ||||| |||
Db    294 C-----CTCTCGTCGTCACATACTACTACACGGTTCTAATCTAATCCGTACTCCACTGCAGAAC 349

Qy    586 GCAGAGCTGGCCACAGAAG--TCCTGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCG 643
          ||||||||||| |||||||  ||||  ||||||| | ||||| |||| ||||||||| |
Db    350 CCAGAGCTGGCCACGCAGCGGCCCAGGAGGCCCGGAATATAGGGCTGACCAGTGACTTTTTCAGTCT 409

Qy    644 GCAAAGGTATGACCCCCTCAGCAGATGTAGTAGTGTAATGTCCCCCTTAGA--TCCCATCCCAGGC 701
          |||||||| || ||  ||||||||||||| |||||   ||||||||||||  |||||||||| |
Db    410 GCGA--GTGTGAGCCCCTCGGCAGATGTGGTAACGCTCCCTGAGACCCCCCCATCCCCCATCCTAGGT 467

Qy    702 AGGTCTCTAAGAGACATGGATGAGAGATGTAGTCATGTGCATTGGCATTCCAAACACAGCTAT 761
          ||||||||||||| ||||| |||||||| || | || || |
Db    468 AGTCTCTCGAGAGGACACAGGAGGATCACACAGATAT-GTCACATGCCATTCCAAACACAGCCAT 526

Qy    762 CCACAGTCGTCCCTTGCCCCTTCCACTTAGCCCAGGAGGACAGTAACCTTAGCCTATCTTTC 821
          ||||| || |||||| |  ||||  ||| |||||  ||||||||||| ||||||| |||||
Db    527 CCACTGTTGCCCATGCTCC--TCCACATAATTAGGAGG--CAGTAACCTTGGCCTATCTTTT 584
```

(Right-hand alignment block)

```
                                                    ...CTTCC 875
                                                    |||||
Db    585 TTCTTTCTCCCACCCTCGTGTCCCACCCAGGACCAACCCCTTGGCTCGTGATACTCAG 644

Qy    876 TTCACGTCCGTCTGTGACTTCCATTTGCAAGGCTTTGACCTCTGCAGCTGCTGGAAGA 935
          ||||||||  |  |||||||||||  |||||||||||||||||||||||||||||||
Db    645 TTCTTCTCTCATTTTCCCTCCATCTGCAAGGCTTTTGACCTCTGCAGCTGCTGGAAGA 704

Qy    936 TAGAGTTTGGCCCTAGGTGTGGCAAGCCATCTCAAGAGAAAGCAGACAACAGGGGGGACCA 995
          |  ||| ||||||||| ||  ||| |||||||||||| |||||| |||| |
Db    705 TGGGATTTGGCTCTTATGCGGCAAAGCCATCTCAAGAGGAGCAGCAGCCACAGAGAACCA 764

Qy    996 GATTTTGGAGGATCAGGAACTAAATCACTGGCGGCCTCGGGGTAGAAAAAGAGTGAG 1055
          |||||||||||||||||||| ||| || |||||||||||||| ||||| | |
Db    765 GATTTTGGAGGAGGAGGACTAAATCATTGAGGGGCCATGAGGGGTAGAAGAGAACAAG 824

Qy   1056 TGAGTCCGTCTAAGCACAAGCTAGTCCCGAGATACTCTGCCACAGCTGGGCTGCT 1115
          |||||||||||||||||| ||||||||||||| || |
Db    825 CAAGTCCTCTGCAGCTAAGCAAGGTAGTTCCCTGCATTCTGCCTCA--GGGCTGCT 881

Qy   1116 CGGGGTAGCTTTAGGAATGTGGGTCTGAAAGACAATGGATTGGAAGACATCTCTT----- 1171
          ||||||| ||||  ||||||  ||||||||| || || |
Db    882 AAGGGCAGGGTTAGGAATGTGGGTTCGAGAGACAATGGGATTTTTAATCCATTCTCTGGA 941

Qy   1172 ------------------TGAGTCTCCCCTCAACAGCCCTGTGTTCTGACCACTGAGCTAGGCAACCAGCA 1272
                            ||||||||| |||||||||| |
Db    942 ACCTGGTTCAGCCCTGGCCCGAATCTCCCCTCAACCATACCTATAGACACATTGTGCAC 1001

Qy   1273 TGGGCCCTGTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAGCAGGGAGGGAG 1332
          ||||||||| || |||||| | |||| | |||  ||| |
Db   1002 GGCCCGACTCCTGTTCAACAACTCTCCCTGGCTCTGATCACTGAGAGCACTGAGAGACCAGAGCA 1061

Qy   1333 AACAGAACGAAATAAGGAAGAAGAAGAAGGCCAGTCAATCAGATGCAGTCAGAAG 1392
          ||||||||| ||   ||||
Db   1062 TGGGCCCTGTACTGAAGTAGAGG---GGTTGCCAACAGCAGCAGCAGGGAGGGGCG 1117

Qy   1393 AGATGGAAGCCAACACACAGCTTGAGCAGGAAAACAGAAAAGGAGAGATTCTGGCA 1452
          ||||| |
Db   1118 AACAGAGAAAGCCCATAAGGAGGAA 1142

Qy   1453 TAAGGAGGCCACAGAAAGAAGAGAGCCCAGGCCCCAAGTCTCCTCTTTTATACCCTCATCC 1512
          ||  || ||||||| |
Db   1143 --AGGAATGTCCTTACTCAGGTTAGACTGGGAGACTGGGGCTGGGGCA 1192

Qy   1513 CGTCTCCCAATTAAGCCCACTCTTCTTCCTCCTAGACTCAGGTGCAGCGAAGGACC 1572
          ||||||||||| ||||||  |||| ||| |
Db   1193 ----GGGCAACACAGAAAGAGAGCCCAAGCCCCCACTTCTTTTACACCCTCAATCCC 1248

Qy   1573 CGTAG------GGAGGATCACACTGGATGAAGAGAAGATAAAGGTGCTTCAAAGGTGCCAGG-CTGTG 1685
          |||||      |
Db   1249 TTCTTCCTCTCTAAACCATTCCTCTCCCTAAGACAGAACTTGAACAGAAGAAGAATC 1308

Qy   1686 CACACAGAGGGTCGAGGACTGGTGTGAGTGAGTGTGTGTAGCCTCAAGATAAGGATGATGTCAGAATGGG 1745
          |||| |
Db   1309 AGTAGTCAGAAGAGGTACACACCGGCCAAGAGGTGATGTGTGGAGAGCCCAGAAAGCT 1368

Qy   1746 CGGGGGGGGGATTCTGGGGGGGGGGAGAGAAGTGAGAGAGAGCCTGGAACAGAGAAT 1805
          |||||||
Db   1369 AAGAGCCAGGACCTAAAAAGACCAAGAGAGCTAAGAGATGCCACCAGGACTACCAGGACTGGG 1428

Qy   1806 CTGGAAGCGCTGGAACGATACCATAAAGGAAGAACCCAGCTACCTTTAGATGTAAAT 1865
          ||||||     | |
Db   1429 CCTGAACAGAGTCAAGGACGCGGTAGAGCCCC----AGCACGAGAGCGTAAGAAATGCA 1483

Qy   1866 CATGAAAGACAGGGAGAAGGGAAGCGGGAGAGTAGAAGAGAACCCGGGCGGGAAGACATGG 1925
          ||||||
Db   1484 CGAG---------GAGAGACAATGGGAAGGGGCCTGGAACGGAGAGAAT 1521

Db   1522 CTGGAAGCCGCTGGAAACGATACCATAAAGGAAGAACCCAGCTACCTTTAGATGTAAAT 1578

Db   1522 CTGGAAGCACAGGGAGTGACAC--TGCAAGGGAAAACTCAGGCTACCTTT-GATGTAAAT 1578
```

Tue Apr 26 16:25:28 2005

us-10-613-728-1.rge

```
Db        ||| |||||                                        || ||||||||||||
          CACGGA------------GCCAAGGAAAAGGAGGAGGCTGGAGAGGAAGAACCCGGGCCAAGACAATGG   1631
Qy  1926  AAGCAAGGACAAGCCAGGTTGAGCGCCTCCGTGGAAATCAGCTCGCTGCTGAAGGCAGAGCCCTG         1985
          |||||||   ||||  |||||||||||||||||| ||| |||||| |||||||||||||||||||
Db  1632  ----------TAGCCAGGGCGAGCACTCGTGTCTGAGATCTGTCTGCAGAA-GCAGAGCCCCTG         1680
Qy  1986  GTATGAGCAGAACAGCAGCAGGCTAGGGTTAATGTCGAGACAGCAGGAACAGAGAGGTAGA           2045
          |||||||||| |||||||||||||||||||||         || ||||||| ||||||||||
Db  1681  GTGCGGAGCGGCCAGAGCAGCAGGAGGTC--------TGAATTCGGCAGCGGACAGAGAGGTAGA       1731
Qy  2046  CACAGGAACAGACAGACG-GGGGAGCCAGGTAACAAGGAAT--------                       2088
          ||||||||||||||||| ||||  ||| ||||||||||
Db  1732  CCCTGGAACAGACAGACCCAAAAACGCAGGAAGCCAGGGAACAAAAGAGCACGGACTGTGCCTCAAAG   1791
Qy  2089  G-------------GGTCCTTCCACCTGTGGCCAGAGCGTCCATCTGTCCACATAC                2133
          |             ||||| ||||||||||||||| |||||| ||||||||||||
Db  1792  CCAGCTTGCCCTTTGGGTCCTTCACCTGTCACCTGTGACCAGAGCGTCCATCTGTCTTCCACATAC    1851
Qy  2134  TCTAGAATGTTCATCAGACTGCAGGGGCTTGGCTTGGAGGCTGGAAAGAGTATGTGAG             2193
          |||||||||||||||||||| ||||||||||||||||  ||||||||||||||||||
Db  1852  TCTAGAACTTTCATCAGACTGCGGGCTG--------GGAGGGCTCGGAAGAGTATGTGAG          1903
Qy  2194  AGCCAGGGGAGACAAGGGGCTTAGGAAAGGAAGGAGAGAAGAAGCAGGCCACACAAGA            2253
          ||||||||||||||||  |||||||||||||||||  ||||||||||||||||||
Db  1904  AGCCAGGAGAGACAAGGGGTTAGGAAAGAAGAAAAAGAAGAAGCCAAACCAGGCCACA---         1960
Qy  2254  GGGCAGAGCCAGAACTGAGTTAACTCCTTCCTTGTTGCATCTTCCATAGGAGGCAGTGG           2313
          |||||||||||||||||||||||||                    |||||||||||||
Db  1961  ----AAAGCCCAGACTGAGTTAACTCTTTCCTTGT------CTCCCACAGGAGACAATGG         2010
Qy  2314  GAACTCTCTGACCACCATCCCATGAGCCCCCACTACCCATCCAACTTTGTCTGCTTGAGT           2373
          |||||||||||||||||||||| |||||| |||| ||||||||||
Db  2011  AGACTCCGTGCCGTGCCCACAATCCCCACAATCCCCACTATACCACCTTTGTCTGCTTGA        2070
Qy  2374  GGCATTCTAGGTTCCCTGAGGACAGAGAGCCTGGCCTTGTCTCTTGGACCTGACCCAAGCT          2433
          |||||| ||||||||||||||||||||| ||||||||||||||||||||||
Db  2071  GGCACACGGGGTTCCCTGAGGACAGAGTATGGTCTTTGTCACTTGGACTTGACCCAGGCT         2130
Qy  2434  GACCCAATGTTCTCAGTAGACCTTTGAATGCCCTCAAGAGTTGAGAACCAGGCAGTGACAT         2493
          ||||||||||||||||||| ||||  |||| |||||||||||||||  ||||||||||||||
Db  2131  GACCCAATGTTCTCAGTAGTCCTTTATCATGTCCCTTATCCTTCAGGAGTTGAAAACCAGGCAGTGACAT   2190
Qy  2494  ATTAGGCACATGGGCTAACCCTGGACTTGCTCCACCACCAGAGCCTCAAGTGACCTCCAGGGAC       2553
          ||||||||  |||| |||||||||| |||||||
Db  2191  ATTAGGCCCACAGGCTAACCCTGTGACTTGCACAAGGTGACCT-----TCCAGGGAC            2241
Qy  2554  ACAGCTCGAGACAGACAGGGCTGGTGCCGCCTTTATCCCCAAAGAGCAACCCATTTGGCATAAGGTCGGCTGCA   2613
          ||||||||||||||||||||||||||||||  |||||||||||||||
Db  2242  CTAGCTCGAGACACAGGTGGCTGCTTGTGTCCATCCTCTGAGAACAATCATTTGGCATAGTCACCTGCA    2301
Qy  2614  AATGGGAATGCAAGGTTGAATCAGTCCCTT----ATACTGCATGCAAGACCTAA                  2669
          ||||||||||||||||||     ||||||| ||  |||||||||| |||||||
Db  2302  GATGGGAATACAAGGTTGGCTCAGTCTGTGTCCGTCAGTGGATAACTAACTAGACTAACATGAGCTAA    2361
Qy  2670  G--------ACCCCTGGAGAGAGGGATGCTCCTGCCCCTGAGCCTGAGGTGCAGGAGGT           2720
          |        ||||| ||||||||| |||||||||||||
Db  2362  GTGTGTGTGTGGCTGGAGAGGGTGTGATGCTGGAGGCTGCTTTTTGGCATGAGCCCCTGAGCCCCAGGAGGT   2421
Qy  2721  GAACTATCCTAGGGGGCCTGGCGCACCTCGCCCCCTGTGTCCTGGGGGAGCAACCACATCTACTGAGTGCTGAGCC  2780
          |||||||| ||||||| |||||||||
Db  2422  GTCCTA-----------GGGGCTGGTGACCTGTGGCTGCCTCCTACCACCTCCACACTCCCACACTGCTCTACC   2474
Qy  2781  CAGAAAACTGACCGCCCCTGTGTCCTGCGCCACCTCCACACTCTTAGAGCTATATTGAGAGG          2840
          |||||||| ||||  |||||||||||||||||||||||||||||||
Db  2475  CAGAGAAATTGACCCGTTGGCCTGGTGTGTGTGTGTTCCTACCTCCACACTCCCACACTCTTGAGAGCTATATTGAGAGG  2534
Qy  2841  TGACAGTAGGATAGGGCGGAGCCAGGAGGTGGGCTAG                                   2900
          ||||||||||  |||||||||||||||||||||||||
Db  2535  TGGCAGTAGAGATGGGGGG-------GTGAGTTCCTGGGGTATGACAGTGTAG                   2578
Qy  2901  GGGAAAGCCAGAGCAGCAGAGGGAGTCTGGCTTTGTCTCTCGAACACAATGTCTACTTACTTAGTTAT    2960
          |||  ||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db  2901  GGGAAAGCCAGAGCAGCAGAGGGGAGTCTGACTTTGTCCCTGAACACAATGTCTCCTTAGTTAT       2960
```

```
Db  2579  GGG-AAGCCAGAGAGGGAGTCTGACTTTGTCCCTGAACACAATGTCTCCTTAGTTAT               2637
Qy  2961  AACAGGCATGACCTGCTAAAGACCCAACATCTACGACCCTGAAAAGACAGCAGCCCTGG             3020
          ||||||||||  ||| |||||||||||| ||||| ||||||||  ||||||||||||||
Db  2638  AACAGGCATGGCCTGTTAAGGACCTAACACCTACGGCCCTCTGAGA--AATGGCTGTCGTA           2695
Qy  3021  AGGACAGGGGTTGTCTCTGAGCCTTGGGTGCTTGATGGTGCCACAAAGGAGGGCATGAGT            3080
          |||||||||   |||| |||| |||||||||||| |||||||||  ||||||||||||
Db  2696  AGGACAGGGAGAACCTCAGTGCCTGGGTGCTTGAAGGTGCCAGAGAATGTTTGTATA               2755
Qy  3081  GTGAGTATAGGCCCAGGACGTTA-GAGAAGGGCACTTGGGAAGGGGTCAGTCTGCAG               3139
          |||||| |||||||||||||||| ||||||||| ||||||||||| ||||||||||||
Db  2756  AGCCCCAGGAGGGGTTAGAGAGGGTACCAGGGTTCTACCTGGGAAGGGGACAGAGTCTGCGG          2815
Qy  3140  AGCCCCTATCCATGGAGCTTGGGGCTCGGGGGCCAACTGGTGTAAATCTCTGG-------            3191
          |||||  ||  ||||||||||||||||||||||||||||||||| |||||||||
Db  2816  AGCCCGCACCAGTGGAATTCGGAGCCTGGGGCCAATGGGGTATAAATCTCTGGAGCTGCCA           2875
Qy  3192  ---GCCTGCCAGGCATTCAAAGCAGCACCTGCATCCTCCTGGGGAGGCGGAA                    3247
             |||||||||||||||||||||||||||||||||||| |||||||||||||
Db  2876  GTCCCCCAGCCAGCCAGCCAGCGGGCCACCTGCATCCTCGGCCTGGAAGGACCGGAA               2935
Qy  3248  GGGAGCAACCCCCCACTTATACCCTTTCTCCCCAGCCCCAGGATTAACACCTCTGGCC-             3306
          |||||||||||| |||||| |||||  ||| |||||||| ||||||||||||||||||
Db  2936  GGGGAGCAACCCCCACTTACACCCCCTCCTCCTCCAGCCCCTAGGATTAACACCTCTGGCCT          2995
Qy  3307  TTTCCCCCTTCCCATCAGGAGTGAGTGAGTGCACGATATATGGACGCAGTGG                    3366
          || ||||| ||| |||||||||||||||||   ||||| |||| ||||||||
Db  2996  TTCCACCCTCTCATCCCCATCAGGAGTGGAGGGTTACAGAGGGGTAAAACCTGTG                 3055
Qy  3367  TGTCCAAACATCATGGTGCACGATATATGGACTCAGTGCCTGAGT                          3426
          |||||||||||||||| ||||| |||||| |||| |||||||||||
Db  3056  TGTCCAAACATGGGGTGCA-GATATGCACCAGTAGACGCAAAAAAGAGAAAT                    3114
Qy  3427  CTGCAGGCTTAACTGGGTTAATGTGTAAAGTC-TGTGTGCATGTGTGTGTCTGACTGA             3485
          ||||| ||||| ||||||||||||||||||| |||||||| ||||||||||||||
Db  3115  CTGCGAGCTTAGCTGGGTTAATGTGTGAAGGCCGTGTGTGTGTGTGTGTGTGT                   3174
Qy  3486  AAACGGGCATGGCCTGTGCAGCTGTGTTCAGTTCTGTGCGTGGTTACCAGGTTT                  3545
          ||||| ||||||| ||||| ||||||||||||| |||||||||||| ||||||||
Db  3175  GTGTGTGGTGTC-GCACAGCTGTTCGGTTGTGTTGTGTGTGAAGTTACTAGACTACAGGTTT          3233
Qy  3546  GTGTGTAAATTGCCCAAGGCAAAGT-GGGTGAATCCTTCCATGGTTAAAGAGATTGGA             3604
          ||| |||||| ||||||||||| ||| ||||||||||||| ||||||| ||||||||||
Db  3234  GTTTGTAAAAATGCCCAAGGCAGAGTGGGGTGAATCCTTCCCGTGGTTTAAAGACATTGGA           3293
Qy  3605  TGATGGCCTGCATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCTCTAA            3664
          |||||   ||||| ||  ||||||| | ||||||||||| || |||||||||||| |||
Db  3294  TGATGG--CGGCATGTCGGGGGACAAT-CAGAGTAAGATAAGATACTCTATATGTGTCCCCAA         3351
Qy  3665  GCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTCTAGAGATGTGGCAACAGAGACTACA           3724
            ||||||||||||| ||||||| ||||||||||||||| || |||||||||||||||||
Db  3352  ACCAAGGTAGCAAGGCCTTTGGACTGTCTCTGGGATACCTGTCTCTAGAGACATGGGCAACAGAGACCATA  3411
Qy  3725  GACAGTATCTGTACAGAGGAGAG-------------AGAGGAGGGGGT                        3763
          |||||  |  |||||||| ||     |||      ||||||||||||
Db  3412  GACAGCATTTATACAGATGAGCTTGGGAGGAGGGGGAGGAGGAGGCAGGAGGGGG                 3471
Qy  3764  GTAGAATTCTCTTACTATCCAAAGGGAAACTGAGTCGTGCACCTGGATGCTCTC                  3823
          |||||||| | ||  | ||||| | |||||||||||||| || ||||||| ||||
Db  3472  AGTAATTTTTTTTTACTAACAAAGAGAAAACTGAGTCATCTGAACAATCTGAACATGATGTCCTC       3531
Qy  3824  CCTAGACATCATGACTTTGTCTCTGGGGAGGGCCTGGGAGCCCAGCACTGTTGGAACTTCAGTCTCAGAGA   3883
          ||| ||||||| |||| | ||  |||||||  ||||| ||||||||||| | |||||| ||||| |||||
Db  3532  CCTTGACATCTTTGACTTTGACTGGGATCTGGGATCCTGGGGCTCTGGGAACTTGGAACTCCAGGTGGAACTGAGAGG  3591
Qy  3884  GTAGGAGGGCTTGAAGTTGTCCAGGGTTGAAAGGGGGGAAGGGA                           3943
          ||||  |  |  | || ||||  |||  || | |||||||| |||
Db  3592  GTAGACGATGCCCCTCCGGGATGGGGTGTGCAGACTGTGGAGGCAGGGGACAGAATTAAGCCTGGAAGAGA-----  3646
Qy  3944  GAGCCTGGGATGGAGCTTGTGTGTTGAGGTTGAAGGCTGGGGATATTAAGCTGTAAGAGA            4003
          |         |||||| ||| || ||| || |||    |||  |||||| ||||||||
Db  3647  --------GGGGAGCCTGTGGGATGGGGGATTCTCTCTGAACTGTCTCTGTAAGAGA               3695
```

RESULT 9
AR054006

```
LOCUS       AR054006                1679 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 11 from patent US 5834306.
ACCESSION   AR054006
VERSION     AR054006.1  GI:5978868
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1679)
  AUTHORS   Webster,K.A. and Bishopric,N.H.
  TITLE     Tissue specific hypoxia regulated therapeutic constructs
  JOURNAL   Patent: US 5834306-A 11 10-NOV-1998;
FEATURES             Location/Qualifiers
     source          1..1679
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN
```

Query Match   22.3%; Score 1279.8; DB 6; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;

```
Qy  3767 GAATTCTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCT 3826
              ||||||||||||||||||||||||||||||||||||| ||||||||||
Db     1 GAATTCTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCT 60

Qy  3827 AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTTCTGAGGAGTA 3886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTTCTGAGGAGTA 120

Qy  3887 GGAGGCTCCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGGAAGGGGAG 3946
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGAGGCTCCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGGAAGGGGAG 180

Qy  3947 CCTGGGATGGGAGCTTGTGTTGGAGGCAGACAGATATTAAGCCTGGAAGGAAGG 4006
              ||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGGGATGGGAGCTTGTGTTGGAGGCAGACAGATATTAAGCCTGGAAGGAAGG 240

Qy  4007 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGGCCT 4066
              |||||||||||||||||||||||||||||||||||||||||||
Db   241 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGGCCT 300

Qy  4067 GGGGTAGGGGAGTCGGTGAGAGACGCTCCGTGTCTCTGTCTCGATGCCCATCGGGCCCTTTG 4126
              ||||||||||||||||||||    |||||||    |||||||||
Db   301 GGGGTAGGGGAGGTCGTGTGTGTCCCAAGACCGTGAGAGACGCCAGGGGCGAGGGCGAA 360

Qy  4127 CAGACCTTTCATGGGCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAAT 4186
              ||||||||||||||||||||||||||||||||||||||||
Db   361 GGGAGGAGGAGGAATGGTGCCCCAAGGACTAAAAAAAAGGCCATGGAGCGGGCGAGGGGCAA 420

Qy  4187 CAGACCTTTCATGGGCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAAT 4246
              |||||||||||
Db   421 CAGACCTTTCATGGGCAAACCTTGGGCCCGTCACCTCACCTGTCTCCTGTCTCCGTGTCTCC 480
```

```
Qy  4004 AGGTGACCCTTACCCAGT-TGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGG 4062
              |||||||||||||||   |||
Db  3696 AGGTGACCCTTACCCCCAGTTGGGGGGAAGGGGAG 3755

Qy  4063 GC--CTGGGTAGGGGAGGTGGTGTGAGACGCTCCGTGTCTCTGTCCATGCCCTGAGGC 4120
              ||  |||||||
Db  3756 CCATGTGGGAAGGGAGGGTGGTGCTATGGTCCTATGTCTTGGCCATCGGGC 3815

Qy  4121 CCTTTGGGGGAGGAGGAATGTGCCCAAGGACTAAAAAAAAAGGCCATGGAGCCAGAGGGGCGA 4180
              ||||
Db  3816 CTTTGGGGGAGGAGGAATGTGCCCAAGGACTAAAAAAAAGGGCCCTGAGCCCAGAGGGGCTA 3875

Qy  4181 GGGCAACAGACCTTTCATGGGCAAACCTTGGGCCC 4216
              |||||
Db  3876 GGGCAACAGACCTTTCATGGGCAAACCTTGGGGCTC 3911

Qy  4247 CGATACCCTTCTTCTTCTAACGGACAGGAGGGAACTCGAGTTCCTATCAGTG 4306
              ||
Db  481 GGATCAAAGGAGGAGGAGCCAGGACAGGAGGGAAGTGGGAGGGAG 525

Qy  4307 ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCATCAGTGATAGAGAAAAGTGAAAGT 4366
              |
Db  526 ------------- 525

Qy  4367 CGAGTTTACCACTCCCATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTA 4426
Db  526 ------------- 525

Qy  4427 TCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCATCAGTGATAGAGAAAAGT 4486
Db  526 ------------- 525

Qy  4487 GAAAGTCGAGTTTACCACTCCCATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4546
Db  526 ------------- 525

Qy  4547 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTCGGTACCAGCAGAGGACTCCAAAT 4606
              |||||||||||||||||
Db  526 ------------GGTCCCAGCAGAGGACTCCAAAT 548

Qy  4607 TTAGGCAGCAGGCATATGGGATGGGGCATATGGGATGGGATATAAAGGGGCACTGAGAGCTGTCAGAG 4666
              ||||||||||||||||||||||||||||||||||||||||
Db  549 TTAGGCAGCAGGCATATGGGATGGGATATGGGATGGGATATAAAGGGGCACTGAGAGCTGTCAGAG 608

Qy  4667 ATTTCTCCAACCCAGTAAGACCGGAGTTCGGGTGGGGGGGCTCTTCACCCACCACCAGACCT 4726
              ||||||||||||||||||||||||||||||||||||||||||||
Db  609 ATTTCTCCAACCCCAGTAAGACCGGAGTTTCGGGTGGGGGGCTCTTCACCCACCACCAGACCT 668

Qy  4727 CTCCCACCTAGAAGAGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT 4786
              ||||||||||||||||||||||||||||||||||||||||
Db  669 CTCCCACCTAGAAGAGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT 728

Qy  4787 GACAGGGTGGCCTTCCACCACCTGGAAGTTCTCAGTGGCAGGAGTTTCCACAAGAAA 4846
              ||||||||||||||||||||||||||||||||||||||||
Db  729 GACAGGGTGGCCTTCCACCACCTGGAAGTTCTCAGTGGCAGGAGTTTCCACAAGAAA 788

Qy  4847 CACTGGATGCCCTTCCCTTACGCTGTCTCCATCTTCCTCTGGGGATGCTCCTCCC 4906
              ||||||||||||||||||||||||||||||||||||||||
Db  789 CACTGGATGCCCTTCCCTTACGCTGTCTCCATCTTCCTCTGGGGATGCTCCTCCCC 848

Qy  4907 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTCGTCCACTCCA 4965
              ||||||||||||||||||||||||||||||||||||||||
Db  849 CGTCTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTCGTCCACTCCA 908

Qy  4967 TCTTTCTCTACTGTCCCCGTGCCCTTGCCCTTGCCTTGCCTTCTCTCCACCCA 5026
              ||||||||||||||||||||||||||||||||||||||||
Db  909 TCTTTCTCTACTGTCCCCGTGCCCTTGCCCTTGCCTTGCCTTCTCTCCACCCA 968

Qy  5027 TTTCTCACTTCACCTTTCTCCCCCTTCCTCCTTCCTTCTGTATTCATCCTTCCTCCTTCCT 5086
              ||||||||||||||||||||||||||||||||||||||||
Db  969 TTTCTCACTTCACCTTTCTCCCCCTTCCTCCTTCCTTCTGTATTCATCCTTCCTCCTTCCT 1028

Qy  5087 TCCTTCCTTCCTTCCTTCCGTGCCGTTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTTATG 5146
              ||||||||||||||||||||||||||||||||||||||||
Db  1029 TCCTTCCTTCCTTCCTTCCGTGCCGTTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTTATG 1088

Qy  5147 TCCTTCCTTCCTTCCGTGTCAGAGTGCTCAGAGTTCAGCAGTCGTGCTGGGGTCTCTTACTTCCTCCA 5206
              ||||||||||||||||||||||||||||||||||||||||
Db  1089 TCCTTCCTTCCTTCCGTGTCAGAGTGCTCAGAGTTCAGCAGTCACACCTGGGGTCTCTTACTTCCTCCA 1148

Qy  5207 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGTCGGGGCAA 5266
              ||||||||||||||||||||||||||||||||||||||||
Db  1149 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGTCGGGGCAA 1208

Qy  5267 CCCCTGGCTTGTCCTGTTCCATCCTGGTCAGATCCTAGATTGGTCTCCAGCCTCTG 5326
              ||||||||||||||||||||||||||||||||||||||||
Db  1209 CCCCTGGCTTGTCCTGTTCCATCCTGGTCAGATCCTAGATTGGTCTCCAGCCCTCTG 1268
```

```
Qy  5327 CTACTCCTTCCTCCTGCCTGTTCCTCCTCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTC 5386
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1269 CTACTCCTTCCTCCTGCCTGTTCCTCCTCTGTCCAGCTGCGCCACTGTGGTGCCTCGTTC 1328

Qy  5387 CAGCTGTGTGTCCACATTCTTCAGATTCTTCAGAAAAGTTAACCAGGTGAGAATGTTTCCC 5446
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1329 CAGCTGTGTGTCCACATTCTTCAGATTCTTCAGAAAAGTTAACCAGGTGAGAATGTTTCCC 1388

Qy  5447 CTGTAGACAGCAGATCACGATTCTCCCGGAAGTCAGCGTTCCAGCCCCTCTCTTCTCTGC 5506
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1389 CTGTAGACAGCAGATCACGATTCTCCCGGAAGTCAGCGTTCCAGCCCCTCTCTTCTCTGC 1448

Qy  5507 CCAGCTGCCCGGCACTCTTAGCAAACTTCAGGCACCCTTACCCCACCATAGACCTCTGACA 5566
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1449 CCAGCTGCCCGGCACTCTTAGCAAACTTCAGGCACCCTTACCCCACCATAGACCTCTGACA 1508

Qy  5567 GAGAAGCAGGCACTTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 5626
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1509 GAGAAGCAGGCACTTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 1568

Qy  5627 GGCAGGGAAGTGGTGGTGTAGGAAAGTCAGGACTTCACATAGAGAGCCTAGCCCACCACCAG 5686
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1569 GGCAGGGAAGTGGTGGTGTAGGAAAGTCAGGACTTCACATAGAGAGCCTAGCCCACCACCAG 1628

Qy  5687 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 5731
         ||||||||||||||||||||||||||||||||||||||||||||
Db  1629 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 1673
```

```
RESULT 10
AR146182
LOCUS       AR146182                1679 bp    DNA     linear       PAT 08-AUG-2001
DEFINITION  Sequence 11 from patent US 6218179.
ACCESSION   AR146182
VERSION     AR146182.1  GI:15109371
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1679)
  AUTHORS   Webster,K.A., Bishopric,N.H., Murphy,B., Laderoute,K.R. and
            Green,C.J.
  TITLE     Tissue specific hypoxia regulated constructs
  JOURNAL   Patent: US 6218179-A 11 17-APR-2001;
FEATURES             Location/Qualifiers
     source          1..1679
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Query Match    22.3%;  Score 1279.8;  DB 6;  Length 1679;
Best Local Similarity 82.2%;  Pred. No. 0;
Matches 1616;  Conservative 0;  Mismatches 57;  Indels 292;  Gaps 1;
```

```
Qy  3767 GAATTCTCTTACTATCAAAGGAAACTGAGTCGTGCACCTGCAAAGTCGATGCTCCTCCCT 3826
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 GAATTCTCTTACTATCAAAGGAAACTCGAGTCGTGCACCTGCAAAGTCGATGCTCCTCCCT 60

Qy  3827 AGACATCATGATGAAGCTGTCTCGGGAGCCAGCACTGTCGAACTTCAGGTCTGAGAGAGTA 3886
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AGACATCATGATGAAGCTGTCTCGGGAGCCAGCACTGTCGAACTTCAGGTCTGAGAGAGTA 120

Qy  3887 GGAGGCTCCCCCAGCCTGAAGCTATGCAGATAGCAGCCAGGGTTGAAAGGGGGAAGGGGAGAG 3946
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GGAGGCTCCCCCAGCCTGAAGCTATGCAGATAGCAGCCAGGGTTGAAAGGGGGAAGGGGAGAG 180

Qy  3947 CCTGGGATGGGAGCTTGTGTTGTTCGGAGGCAGGGGACAGATATTAAGCCTGGAAGAGAAGG 4006
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 CCTGGGATGGGAGCTTGTGTTGTTCGGAGGCAGGGGACAGATATTAAGCCTGGAAGAGAAGG 240

Qy  4007 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGGCCT 4066
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   241 TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGGCCT 300
```

```
Qy  4067 GGGTAGGGGAGGTGGTGTGAGACGCTCCTGTCTCTCCTCTGCATGCCCTGAGGCCCTTTG 4126
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   301 GGGTAGGGGAGGTGGTGTGAGACGCTCCTGTCTCTCCTCTGCATGCCCTGAGGCCCTTTG 360

Qy  4127 GGGGAGGAGGAATGTGCCCAAGGACTAAAAAAAGGCCATGAGAGGGGCGAGGGCGAGGGCAA 4186
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   361 GGGGAGGAGGAATGTGCCCAAGGACTAAAAAAAGGCCATGAGAGGGGCGAGGGCGAGGGCAA 420

Qy  4187 CAGACCTTTCATGGGCAAACCTTGGGCCCTAGTGATCGATTGACAAGAACTCGCCAAT 4246
         ||||||||||||||||||||||||||||||||||        | |||||||||
Db   421 CAGACCTTTCATGGGCAAACCTTGGGCCCTGCTGTCCTGTCACCTCCAGAGCCAAG 480

Qy  4247 CGATACCCTTCTTCTTCTAACGACAGGAGGGACTCGAGTTTACCACTCCCTATCCAGTG 4306
         |||||||||||||||||||||||||||||||||||||||||||||        ------
Db   481 GGATCAAAGGAGGAGCGAGGCAGGACAGGAGGAGGTGGGAGGGAG--------------- 525

Qy  4307 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4366
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 ------------------------------------------------------------ 525

Qy  4367 CGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCCCTA 4426
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 ------------------------------------------------------------ 525

Qy  4427 TCAGTGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGT 4486
         |||||||||||||||||||||||||||||||||||||||||||||||
Db   526 ------------------------------------------------------------ 525

Qy  4487 GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4546
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   526 ------------------------------------------------------------ 525

Qy  4547 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTCGGTACCAGCGAGGAGGACTCCAAAT 4606
         |||||||||||||||||||||||||||||||||||||||||||||
Db   526 --------------GGTCCCAGCGAGGAGACTGAGAGCTGTCGAGAGCACTGAGAGCTGTCAGAG 4666

Qy  4607 TTAGGCAGCAGGCATATGGGATGGGATGGGATATAAAGGGCTGAGAGCACTGAGAGCTGTCAGAG 4666
         |||||||||||||||||||||||||||||||||||||||||||||
Db   549 TTAGGCAGCAGGCATATGGGATGGGATGGGATATAAAGGGCTGAGAGCACTGAGAGCTGTCAGAG 608

Qy  4667 ATTTCTCCAACCCAGGTAAGAGGGAGTTTCGGTGGGGGCTCTTCACCCACCCACCAGACCT 4726
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   609 ATTTCTCCAACCCAGGTAAGAGGGAGTTTCGGGTGGGGGCTCTTCACCCACCCACCAGACCT 668

Qy  4727 CTCCCACCTAGAAGGAAACTGCCTTTCCTGGGAAGTTCAGGCCGGTTCAGGCCGGTCAGAGATCT 4786
         |||||||||||||||||||||||||||||||||||||||||||||
Db   669 CTCCCACCTAGAAGGAAACTGCCTTTCCTGGGAAGTTCAGGCCGGTTCAGGCCGGTCAGAGATCT 728

Qy  4787 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAAA 4846
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   729 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAAA 788

Qy  4847 CACTGGATGCCCCTTCCTTAGCGTCTCCATCTTCCTCCTGGGGATGCTCCTCCC 4906
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   789 CACTGGATGCCCCTTCCTTAGCGTCTCCATCTTCCTCCTGGGGATGCTCCTCCC 848

Qy  4907 CGTCTTGGTTTATCTTGGCTCTTCGTCTTGTCTGCTCGTCGTCCACTCCA 4966
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   849 CGTCTTGGTTTATCTTGGCTCTTCGTCTTGTCTGCTCGTCGTCCACTCCA 908

Qy  4967 TCTTTCTCTACTGTCTCCGTGCCTTCTTGCGTTCCTTGCGTGTCCTTTTCCACCCA 5026
         |||||||||||||||||||||||||||||||||||||||||||||
Db   909 TCTTTCTCTACTGTCTCCGTGCCTTCTTGCGTTCCTTGCGTGTCCTTTTCCACCCA 968

Qy  5027 TTTCTCACTTCACCTTTCTCCTTCCTTTCTCATTTGTATTGTATTCATCCTTCCTTCCT 5086
         ||||||||||||||||||||||||||||||||||||||||||||||
Db   969 TTTCTCACTTCACCTTTCTCCTTCCTTTCTCATTTGTATTGTATTCATCCTTCCTTCCT 1028

Qy  5087 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 5146
         ||||||||||||||||||||||||||||||||||||||||||||||
Db  1029 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 1088
```

```
Qy    5147  TCCTTCCTTCCTTCCTTCCTCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCCTTATG  5205
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1089  TCCTTCCTTCCTTCCTTCCTCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCCTTATG  1148

Qy    5207  TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTTGCTGCTGGGGTGCTCTGTTACCTTCCTCA  5266
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1149  TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTTGCTGCTGGGGTGCTCTGTTACCTTCCTCA  1208

Qy    5267  CCCCCTGGCTTGTCCGTGTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG  5326
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1209  CCCCCTGGCTTGTCCGTGTCCATCCTGGTCAGGATCTCTAGATTGTCTCCCAGCCTCTG  1268

Qy    5327  CTACTCCTCTTCTGCCTGTTCCTCTCTTGTCCGCAGCTGCCCACTGTGCCTCGTTC  5386
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1269  CTACTCCTCTTCTGCCTGTTCCTCTCTTGTCCGCAGCTGCCCACTGTGCCTCGTTC  1328

Qy    5387  CAGCTGTGTCCACATTCTTCAGGAATTCTTGAAAAGTTAACCAGTGAGAATGTTTCC  5446
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1329  CAGCTGTGTCCACATTCTTCAGGAATTCTTGAAAAGTTAACCAGTGAGAATGTTTCCC  1388

Qy    5447  CTGTAGACAGCAGATCACGAATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTTCTTTCTGC  5506
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1389  CTGTAGACAGCAGATCACGAATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTTCTTTCTGC  1448

Qy    5507  CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTGACA  5566
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1449  CCAGCTGCCCGGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTGACA  1508

Qy    5567  GAGAAGCAGGCAGCCACTTTACATGGAAGTCCTGGTGTGTAGGAAGTCAGGACTTCACATAGAAAGA  5626
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1509  GAGAAGCAGGCAGCCACTTTACATGGAAGTCCTGGTGTGTAGGAAGTCAGGACTTCACATAGAAAGA  1568

Qy    5627  GGCAGGGAAGTCGTGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCACCAG  5686
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1569  GGCAGGGAAGTCGTGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCACCAG  1628

Qy    5687  AAATGACAGACAGATCCCCTACATCCCCCATCCCCCATAAGAGTTTGAGT  5731
                  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1629  AAATGACAGACAGATCCCCTACATCCCCCATCCCCCATAAGAGTTTGAGT  1673
```

RESULT 11
AX546496
LOCUS   AX546496   1679 bp   DNA   linear   PAT 26-NOV-2002
DEFINITION Sequence 5 from Patent WO0249669.
ACCESSION   AX546496
VERSION   AX546496.1 GI:25811663
KEYWORDS   .
SOURCE   Mus sp.
  ORGANISM Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Schu,C.A.; Padua,R.; Bonner,M.; Donovan,M.G. and Soykan,O.
  TITLE   Electrically responsive promoter system
  JOURNAL   Patent: WO 0249669-A 5 27-JUN-2002;
            Medtronic, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..1679
                     /organism="Mus sp."
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:10095"
ORIGIN

```
Query Match       22.3%;  Score 1279.8;  DB 6;  Length 1679;
Best Local Similarity  82.2%;  Pred. No. 0;
Matches 1616;  Conservative 0;  Mismatches 57;  Indels 292;  Gaps 1;

Qy    3767  GAATTCTTACTATCAAAGGGAAACCTGAGTCGTGCACCTGCAAAGTGGACGTCTCCCCT  3826
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1  GAATTCTTACTATCAAAGGGAAACCTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCCT  60

Qy    3827  AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA  3886
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      61  AGACATCATGACTTTGTCTCTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA  120

Qy    3887  GGAGGCTCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTGAAAGGGGGAAGGGGAGAG  3946
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121  GGAGGCTCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTGAAAGGGGGAAGGGGAGAG  180

Qy    3947  CCTGGATGGGAGCTTGTGTGTTGGAGGCAGGGGACAGATATTAAGCTGGAGAGAGAAGG  4006
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181  CCTGGATGGGAGCTTGTGTGTTGGAGGCAGGGGACAGATATTAAGCTGGAGAGAGAAGG  240

Qy    4007  TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATAAAAATAACTAACTGAGGTAAGGGCCT  4066
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241  TGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATAAAAATAACTAACTGAGGTAAGGGCCT  300

Qy    4067  GGGTAGGGAGGGAGGTGGTGTGAGACGCTCCTCCTGTCTCCTCCTCTGCCATGCGATGCCCTGTCCCATG... GAGGCCCTTTG  4126
                  |||||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db     301  GGGTAGGGAGGGAGGTGGTGTGAGACGCTC---CTGTCTCCTCCTCTGCCATGCGGTGCCCTTG... GAGCCCCTTTG  360

Qy    4127  GGGAGGAGGAATGTGCCCAAGACTAAAAAAGGCCATGAGCCAGGGGGCAGAGGGCGGAGGGCAA  4186
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361  GGGAGGAGGAATGTGCCCAAGACTAAAAAAAAGGCCATGAGCCAGGGGGCAGAGGGCGGAGGGCAA  420

Qy    4187  CAGACCTTTCATGGGCAAACCTTGGGCGCCCGTAGTGATGCATTGACAAGAACTCGCCAAT  4246
                  |||||||||||||||||||||||||||||||||||||||    |||||||||||||||||||
Db     421  CAGACCTTTCATGGGCAAACCTTGGGCGCCCGTCGTCCTGTCCTGTCACCTCCAGACCCAAG  480

Qy    4247  CGATACCCTTCTTCTTCTAACGGACAGGAGGGAACTCGAGTTTAAAGA  4306
                  ||||||||||||||||||    |||||||||||||||||||||||
Db     481  GGATCAAAGGAGGAGAGCCAGGAGCCAGGAGAGTGGGAGGAG  525

Qy    4307  ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGT  4366
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4307  ATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGT  4366

Qy    4367  CGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTA  4426
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     526  ------------------------------------------------------------  525

Qy    4427  TCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGT  4486
                  ------------------------------------------------------------
Db     526  ------------------------------------------------------------  525

Qy    4487  GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC  4546
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     526  -----------------------------------------------------------  525

Qy    4547  TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTCGAGTCCAGCAGAGACTCCAAAT  4606
                         |||   ||||||||||||||||||||||||||||||||
Db     526  ---------------------GGTCCCAGCAGAGACTCCAAAT  548

Qy    4607  TTAGGCAGCAGCATATGGGATGGGAATGGGAGACTGTGGAGCACTGGAGAGTGTCAGAG  4666
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     549  TTAGGCAGCAGCATATGGGATGGGAATATAAAGGGGATAATAAAGGGGAGACTGGAGAGTGTCAGAG  608

Qy    4667  ATTTCTCCAACCCAGTGAAGGGAAACTGCTTTCACCACCAGACCTCCACCACCAGACCT  4726
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     609  ATTTCTCCAACCCAGTGAAGGGAAACTGCTTTCACCACCAGACCTCCACCACCAGACCT  668

Qy    4727  CTCCCACCTAGAAGGAAGAAACTGCCTTTCCTGGAAGTTCAGGCGGTTCAGGCGAGATCCAG  4786
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     669  CTCCCACCTAGAAGGAAGAAACTGCCTTTCCTGGAAGTGGGTCAGGCGGTTCAGGCGAGATCT  728

Qy    4787  GACAGGTGGCCTTCCACCAGCTGGCAAGTTCTCAGTGGCAGGAGTTTCCACAAGAAA  4846
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     729  GACAGGTGGCCTTCCACCAGCCTGGCAAGTTCTCAGTGGCAGGAGTTTTCCACAAGAAA  788

Qy    4847  CACTGGATGCCCCTTCCCTTCGTCTTCCTTACGCTGCTGATCTTCTTCCATCTTCTCCATCTTCTCCTCCC  4906
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     789  CACTGGATGCCCCTTCCCTTCGTCTTCCTTACGCTGCTGATGGGATGCTCCTCCATCTTCTCCCT  848

Qy    4907  CGTCTTGGTTTATCTTGCCTCTTCGTCTTCAGCAAGATTGCCCGTGCGTCGTCCACTCCA  4966
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
COMMENT        Original source text: Mus musculus DNA.
FEATURES             Location/Qualifiers
     source          1..2594
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
     mRNA            join(603..623,1305..1373,1665..1880)
                     /product="cardiac myosin heavy chain"
     exon            603..623
                     /number=1
     intron          624..1304
                     /number=1
     exon            1305..1373
                     /number=2
     intron          1374..1664
                     /number=2
     exon            1665..1880
                     /number=3
     CDS             1680..1880
                     /codon_start=1
                     /product="cardiac myosin heavy chain"
                     /protein_id="AAA37424.1"
                     /db_xref="GI:192610"
                     /translation="MTDAQMADFGAAAQYLRKSEKERLEAQTRPFDIRTECFVPDDKE
                     EYVKAKVVSREGGKVTAETENGK"
     intron          1881..>2594
                     /number=3
ORIGIN

Query Match             22.3%;  Score 1279.8;  DB 10;  Length 2594;
  Best Local Similarity   82.2%;  Pred. No. 0;
  Matches 1616;  Conservative    0;  Mismatches   57;  Indels  292;  Gaps  1;

Qy   3767 GAATTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGAAAGTGGATGCTCTCCCT 3826
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GAATTCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGAAAGTGGATGCTCTCCCT   60

Qy   3827 AGACATCATGACTTTGTCTCTGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA 3886
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGACATCATGACTTTGTCTCTGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTA  120

Qy   3887 GGAGGCTCCCCTCAGCCTGAAGCTATGCAAGCCTGCAGATAGCAGGGGGAAGGGAAGGGAGAG 3946
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GGAGGCTCCCCTCAGCCTGAAGCTATGCAAGCCTGCAGATAGCAGGGGGAAGGGAAGGGAGAG  180

Qy   3947 CCTGGGATGGGAGCTTGTGTTGTGGAGGCAGGGACAGATATTAAGCCTGGAAGAGAGAAGG 4006
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CCTGGGATGGGAGCTTGTGTTGTGGAGGCAGGGACAGATATTAAGCCTGGAAGAGAGAAGG  240

Qy   4007 TGACCCTTACCCAGTGTTCAACTCACCCTTCAGATTAAAATAACTAACTGAGGTAAGGGCCT 4066
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TGACCCTTACCCAGTGTTCAACTCACCCTTCAGATTAAAATAACTAACTGAGGTAAGGGCCT  300

Qy   4067 GGGTAGGGGAGGGTGGTGTGGAGACGGCTCCTGTCTCCTCGTCTCCTCCTCCTCGCATGCCCATCGGCCATCGGCCCTTTG 4126
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGGTAGGGGAGGGTGGTGTGGAGACGGCTCCTGTCTCCTCGTCTCCTCCTCCTCGCATGCCCATCGGCCATCGGCCCTTTG  360

Qy   4127 GGGAGGAGGAATGTGCCAAGGACTAAAAAAGGCCATGAGCCCAGAGGGGGCGAGGGCAA 4186
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGGAGGAGGAATGTGCCAAGGACTAAAAAAGGCCATGAGCCCAGAGGGGGCGAGGGCAA  420

Qy   4187 CAGACCTTTCATGGGCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAAT 4246
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CAGACCTTTCATGGGCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAAG  480

Qy   4247 CGATACCCTTCTTCTTCTAACGACAGGAGGACGAGGAACTCGAGTTTACCACTCCCTATCCAGTG 4306
          |||||||||||                   ||||||||||||||||||||||||||||||||||
Db    481 GGATCAAAGGAGGAGGAGCCAGGACAGGAGGAAGTGGGAGGGAG------------------     525

Qy   4307 ATAGAGAAAAAGTGAAAGTCGAGTTTACCACTCCCTATCCAGTGATAGAGAAAAAGTGAAAGT 4366
          ----------------
Db    526 ------     525
```

RESULT 12
MUSCMHCG

```
LOCUS       MUSCMHCG                2594 bp    DNA     linear   ROD 27-APR-1993
DEFINITION  Mus musculus cardiac myosin heavy chain gene, 5' end.
ACCESSION   M62404
VERSION     M62404.1  GI:192609
KEYWORDS    cardiac myosin heavy chain.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 2594)
  AUTHORS   Gulick,J., Subramaniam,A., Neumann,J. and Robbins,J.
  TITLE     Isolation and characterization of the mouse cardiac myosin heavy
            chain genes
  JOURNAL   J. Biol. Chem. 266 (14), 9180-9185 (1991)
  MEDLINE   91225025
   PUBMED   2026617
```

```
Db    849 CGTCTTGGTTATCTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGCTGTCCACTCCA  908
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   4967 TCTTTCTCACTGTCTCCGCGCCTTGCCTTCGGCCTTGCCGTGCTGCCTTCCTTCCTTCCACCCA 5026
          ||||||||||||||||||  |||||||||||||||||||||||||||||||||||||||||||
Db    909 TCTTTCTCACTGTCTCCGTGCCTTGCCTTCGGCCTTGCCGTGCTGCCTTCCTTCCTTCCACCCA  968

Qy   5027 TTTCTCACTTTCACCTTTCTCCCCTTCTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTT 5086
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    969 TTTCTCACTTTCACCTTTCTCCCCTTCTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTT 1028

Qy   5087 TCCTTCCTTCCTTTCTCCCCTTCTCCTTCCTTCCT 5146
          |||||||||||||||||||||||||||||||||||
Db   1029 TCCTTCCTTCCTTTCTCCCCTTCTCCTTCCTTCCT 1088

Qy   5147 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACACCCTGGGGTTCCCACCCTTATG 5206
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1089 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACACCCTGGGGTTCCCACCCTTATG 1148

Qy   5207 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTCTCCTCA 5266
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1149 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTCTCCTCA 1208

Qy   5267 CCCCTGGCTGTCTCGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 5326
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1209 CCCCTGGCTGTCTCGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG 1268

Qy   5327 CTACTCCCTCCTCCTGCCTGTTCCTCCTCTGTCCAGCTGCCGCCACTGTGGTGCCCGTTC 5386
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1269 CTACTCCCTCCTCCTGCCTGTTCCTCCTCTGTCCAGCTGCCGCCACTGTGGTGCCCGTTC 1328

Qy   5387 CAGCTGTGGTCCACATTCTTCAGGATTCTTGAGAAAGTTAACCAGTGAGAATGTTCCC 5446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1329 CAGCTGTGGTCCACATTCTTCAGGATTCTTGAGAAAGTTAACCAGTGAGAATGTTCCC 1388

Qy   5447 CTGTAGACAGGAGATCACGATTCTCCGGAAGTCAGGCTTCCAGCCCTTCTTTCTTCTGC 5506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1389 CTGTAGACAGGAGATCACGATTCTCCGGAAGTCAGGCTTCCAGCCCTTCTTTCTTCTGC 1448

Qy   5507 CCAGCTGCCCGGCCACTCTTAGCAAACCTCAGGCCACCCCACCCTACCCCACATAGACCTCGACA 5566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1449 CCAGCTGCCCGGCCACTCTTAGCAAACCTCAGGCCACCCCACCCTACCCCACATAGACCTCGACA 1508

Qy   5567 GAGAAGCAGGCAGTGGTGTGGTGTAGGACTTCACATGGAAGTCAGGAAGTTGAGGATAGCCATAGGCTACGGTGTAAAAGA 5626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1509 GAGAAGCAGGCAGTGGTGTGGTGTAGGACTTCACATGGAAGTCAGGAAGTTGAGGATAGCCATAGGCTACGGTGTAAAAGA 1568

Qy   5627 GGCAGGGAAGTGGTGTGTGTGTGGTGTGTGGGCAGGGACTTCACATAGAAGCCTAGCCCCACCACCAG 5686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1569 GGCAGGGAAGTGGTGTGTGTGTGGTGTGTGGGCAGGGACTTCACATAGAAGCCTAGCCCCACCACCAG 1628

Qy   5687 AAATGACAGACAGATCCCCTCCTATCTCCCCATAGAAGTTTGAGT 5731
          |||||||||||||||||||||||||||||||||||||||||||||
Db   1629 AAATGACAGACAGATCCCCTCCTATCTCCCCATAGAAGTTTGAGT 1673
```

```
Qy   4367 CGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCTA 4426
                                                                                Db   1389 CTGTAGACAGCAGATCACGACGATTCTCCCGGAAGTCAGGCTTCCACCCTCTTTCTCTGC 1448
Db    526 ------------------------------------------------------------- 525
                                                                                Qy   5507 CCAGCTGCCCGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTCTGACA 5566
Qy   4427 TCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGT 4486
                                                                                Db   1449 CCAGCTGCCCGCACTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTCTGACA 1508
Db    526 ------------------------------------------------------------- 525
                                                                                Qy   5567 GAGAAGCAGGCACTTTACATGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 5626
Qy   4487 GAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4546
                                                                                Db   1509 GAGAAGCAGGCACTTTACATGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGA 1568
Db    526 ------------------------------------------------------------- 525
                                                                                Qy   5627 GGCAGGGAAGTGGTGGTGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACCAG 5686
Qy   4547 TCCCTATCAGTGATGAGAGAAAAGTGAAAGTCGAGTACCAGCAGAGGACTCCAAAT 4606
                                                                                Db   1569 GGCAGGGAAGTGGTGGTGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACCAG 1628
Db    526 ---------------------------GGTCCCAGCAGAGGACTCCAAAT 548
                                                                                Qy   5687 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 5731
Qy   4607 TTAGGCAGGCAGCATATGGGATGGGATATAAAGGGGCTGGAGCACTGAGACGTGTCAGAG 4666
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    549 TTAGGCAGGCAGCATATGGGATGGGATATAAAGGGGCTGGAGCACTGAGACGTGTCAGAG 608
                                                                                Db   1629 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 1673

Qy   4667 ATTTCTCCAACCCAGGTAAGAGAGAGTTCGGGTGGGGGCTCTTCACCCACCAGAGACCT 4726
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    609 ATTTCTCCAACCCAGGTAAGAGAGAGTTTCGGGTGGGGGCTCTTCACCCACCAGAGACCT 668

Qy   4727 CTCCCCACCTAGAAGGAAAACTGCCTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT 4786
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    669 CTCCCCACCTAGAAGGAAAACTGCCTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCT 728

Qy   4787 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAAA 4846
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    729 GACAGGGTGGCCTTCCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAAA 788

Qy   4847 CACTGGATGCCCCTTCCCTACGCTGTCTCTCCATCTTCCTCCTGGGGATGCTCCTCCC 4906
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    789 CACTGGATGCCCCTTCCCTCCGTCTTCCTCCATCTTCCTCCTGGGGATGCTCCTCCC 848

Qy   4907 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGTCCACTCCA 4966
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    849 CGTCTTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGTCCACTCCA 908

Qy   4967 TCTTTTCTCTACTGTCTCCGGTCCCTGCCTTGCCCTTGCGCGTGTCCTTCCTTCTCCACCCA 5026
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    909 TCTTTTCTACTGTCTCCGGTCCCTGCCTTGCCCTTGCGCGTGTCCTTCCTTCTCCACCCA 968

Qy   5027 TTTCACTTCACCTTTCTCCCCCTTCTCCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCT 5086
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    969 TTTCACTTCACCTTTCTCCCCCTTCTCCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCT 1028

Qy   5087 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 5146
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1029 TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 1088

Qy   5147 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATG 5206
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1089 TCCTTCCTTCCTTCCTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATG 1148

Qy   5207 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGCTGGGGTGCTCTCTAACCTTCCTCA 5266
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1149 TAAACAATCTTCCAGTGAGCCACAGCTTCAGTGCTGCTGGGGTGCTCTCTAACCTTCCTCA 1208

Qy   5267 CCCCCTGGCTTGTCCTGTTCCATCCTGGTCAGGATCCTCTAGATTGGTCTCCCAGCCTCTG 5326
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1209 CCCCCTGGCTTGTCCTGTTCCATCCTGGTCAGGATCCTCTAGATTGGTCTCCCAGCCTCTG 1268

Qy   5327 CTACTCCTCCTCCTGTTCCTCCTTCCTGCAGCTGCGCCACTGTGGCTGTGCTGCTCGTTC 5386
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1269 CTACTCCTCCTCCTGTTCCTCGTCCTCCTGCAGCTGCGCCACTGTGGCTGTGCTGCTCGTTC 1328

Qy   5387 CAGCTGTGGTCCACATTCTTCAGGATTCTCTGAAAAGTTAACCAGTGAGAATGTTTCCC 5446
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1329 CAGCTGTGGTCCACATTCTTCAGGATTCTCTGAAAAGTTAACCAGTGAGAATGTTTCCC 1388

Qy   5447 CTGTAGACAGCAGATCACGACGATTCTCCCGGAAGTCAGGCTTCCACCCTCTTTCTCTGC 5506
```

```
FEATURES        Location/Qualifiers
source          1..157910
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="14"
                /clone="C-2201G16"
                /clone_lib="CalTech-D"
STS             13550..13694
                /note="matching EMBL:AA452257
                RHdb:RH92419
                dbSTS:STS65422
                Identified using the e-PCR software (G. Schuler)"
STS             46758..46857
                /note="matching EMBL:N91549
                RHdb:RH76320
                dbSTS:STS53407
                Identified using the e-PCR software (G. Schuler)"
STS             57409..57527
                /note="matching EMBL:X52889
                RHdb:RH96140
                RHdb:RH13799
                dbSTS:STS7881
                Identified using the e-PCR software (G. Schuler)"

ORIGIN

Percentage of bases with a quality value >= 40 : 99 %.

Query Match          9.7%;   Score 558.2;   DB 9;   Length 157910;
Best Local Similarity  63.7%;   Pred. No. 3.4e-164;
Matches 1380;   Conservative 0;   Mismatches 658;   Indels 127;   Gaps 30;

Qy   2096  CTCACCTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGC 2155
Db  44522  CTCACCTGTGGCCGAAGCCTCCATCCCTGTCCTTATATTCCAGACTCCCAGCAGAATGA 44463

Qy   2156  AGGGCT------GGCTTGGGAGCAGCTGGAAAGAGTATGTGAGAGCCAGGGGAGACA 2207
Db  44462  GGAGCTGTGTGCAAGGCAGCTTAGGGAGGACGGAGCAATATGTGAGAGCAGGAGAGA 44403

Qy   2208  AGGGGGCCTAGGAAGAAGAGAGGAGAAGAGGGCAAACCAGGCCACACAAGAGAGGCCAGA 2267
Db  44402  A-GGAGCCTGGGGAGGAGGGAGGACGGTACAGCCGAGCAGAC----GGGCGAGGGCCTGGG 44348

Qy   2268  ACTGAGTTAACTCCTTCGTTGTTGCATCTTCCATAGGAGGCAGTGGAACTCTGTGACCA 2327
Db  44347  GCTGAGTAGTGAACACCTTCCTTCCTCCTCCTTCCCACAGGAGGGTAAAAACCTACATGTGTCCAAACATCATGG 44288

Qy   2328  CCATCCCCATG------AGCCCCCACTACCCATCCCAAGTTTGCCTGAGTGGCATTCT 2381
Db  44287  CTCCACTCCCCGCTGCCCGTCCCCACCACCATATCCACCTTAGAGGCCTGCTGCTCAGAGCACA 44228

Qy   2382  AGGTTCCCTGAGGACAGAGCCTGGCCTTTGTCTCTTGGACCTGACCCCAGCTGACCCCAAT 2441
Db  44227  GGGTTCCCTGAGGACACAAAGCCTGGTCTTTGTCTGTCACCTGACCCCGACCCCAGACCTGGC 44168

Qy   2442  GTTCTTCAGTACCTTTGAATGCCCTCAAGAGCTTGAGAGAGTTGAGCACATATTAGGCC 2501
Db  44167  GTTCTTCAGCTCCTTATCATGTCCCCAGAGCTATGGGAACCAAGCAGTCAGTGTGTAAATTGCC 44109

Qy   2502  ATGGGCTAACCCTGAGCTTGCACACAGAGCCTCAAGTGCACCTCCAGGACACAGCCTGC 2561
Db  44056  GGACAGGTGCCTGCATCCCCCTGGGGTAATGATTTGGCATAAGCACCTGCAGGAGAATGG 43997

Qy   2562  AGACAGGTGGCCTTTATCCCCAAAGAGCAACCATTTGGCATAGTGGTCGGCTGC---AAATGG 2618
Db  43996  GGAATGTGGGGTCAACTAAGGTCCCTCCAAGGACATTG-AGAAAAGGTCCAAATGCTCTGG 43938

Qy   2619  GAATGCAAGGTTGAATCAGGTCCCTTCAAGAATACTGCATGCAAGACACCCCTGG 2678
Db  43937  AAAGAGAGGTGTGTTCCTGAGCCCA------GGAATGGGGTGGCAGCGTCCTAGGGAGCT 43884

Qy   2679  AGAGAGGGGGTATGCTCCTGCCCCACCCACCATAAGGGAGTGAACTATCCTAGGGGGCT 2738
Db  43883  AGGG------GAGAGGCCACATCTTAATGAAAGTGCTCCTACCCAGAGAAGCTGACCCCT 43829

Qy   2739  GGCGACCTTGGGAGACACCACATTACTGAGAGTGCTGAGCCAGAAAACTGACCGCCC 2798
Db  43828  GTCGTCCTACCCACCTCCACACCCTCGAGCTATATTGAGAGGTGACAGTAAACTGGTGGG 43769

Qy   2799  TGTGTCCTGCCCACCTCCACACTCTTAGACTCTTCGAGGTGACAGTATATTGAGAGGTGG 2858
Db  43768  GAGCTGGTAGCAGGGGAGGAGGAGTGTTCCTGGGTGTGAGGGTGTAGGGGTGACAGAGG 43713

Qy   2859  ---AGGGGGCTGGGAGGGTGTTCCTGGGTGTGAGGGGTG--GGGGGAAGGCCAGAGCAGGG 2918
Db  43712  GAGTCTGGCTTTGTTGTTCCTGAACACAATGTCTACTTAGTTATAACAGGCATGACCTGCTA 43653

Qy   2919  GAGTCTGGCTTTGTTGTTCCTGAACACAATGTCCACTTAGTCATAACAGGAACAGCCTGCTG 2978
Db  43652  AAGACCTGACACTTACAGCCTCTG-AGAGATGGCAGCATTGGAGGATGGGGAGCGGGGCC 43594

Qy   2979  AAGACCCACACATCTACGACCTCTGAAAAGACAGCAGCCCTGGAGA-----CAGGGGTTGT 3034
Db  43593  TTGAGGGGCCTAGGGTGCTTGAGGACGTCCAAAGGGCCAAAGAGAGG---GAATGTGAGTATAAGCCC 43537

Qy   3035  CTCTGAGCCTTGGGTGCTTGATGTGGTCCACAAAGAGAGGGCATGGGGCCATGCGTGTGAGTGTGAGTATAAGGGCC 3094
Db  43536  TGGAGGCCTTGGAAAGGGGCAACAAAATTCTGCCTGGGGAGGGGCAGGGTGCAGAGCCC 43477

Qy   3095  CCAGGAGC--------GTTAGAGAAGGGCACTTGGGAGGGGTCAGTCTGCAGAGCCC 3144
Db  43476  CTATCC-ATGGAATCTGGAGCCTGGGG-CCAACTCGTGTAAATCTTCTGGCCTGCCAGGC 43419

Qy   3145  CTATCC-ATGGAATCTGGAGCCTGGGG-CCAACTCGTGTAAATCTTCTGGCCTGCCAGG-- 3202
Db  43418  CACCCCGATGGAATGCAGTGCCTGGGACTCCCAGCCTGTAAATCCCCAGGCCTGCCAG-- 43369

Qy   3203  ATTCAAAGACAGCATCCATCCCTCTGCGAAGCGTGGGGCGGAGGGAGCAACCCCCA 3262
Db  43368  ------GCACCTGCACCCCTCGGCAGCCCGGTGAGCGCAGAAGCAGCAGCCCCTCA 43313

Qy   3263  CTTTATACCCTTCTCCCTCAGCCCCAGGATTAACACCTCTGGCCTTCCCACCT 3322
Db  43313  CCCACCCCCTCCTCTCTCAGGCCTCTCAGGGATTAGCACTCTTCTT----CTGCCCACCCC 43313

Qy   3323  CCCATCAGGAGTGGAGGGTTGCAGAGGGCAGAGGGTAAAAAACTACATGTCCAAACATCATGG 3382
Db  43312  CACCCCAGGAGCAGAGGGGTAGCAG--GGGAGGGCAAAAAGCCTATGTGTGCAAATGCGGATG 43254

Qy   3383  TGCACGATATATGGAATCAGTCAGTGTGTAGAGGCA---AGAAAGGAAAATCTGAGGCTTAAC 3439
Db  43253  TGCA--AATGTGTCATCCATGTACACAAAGAAGTAGAAACAGAAAAGAAATCTGCAAGCTTGAG 43195

Qy   3440  TGGGTTAATGTGTAAAGTCTGTGTGCATGTGTCTGTGAAGAAAATCTGGCTGTGTGTGAAAATTGCC 3499
Db  43194  TGGGTAAGTGTGAAAAATCTGCATGTGTGGCTGGCAGGAGAATGGGCAACAGACACCG-------- 43143

Qy   3500  GTGCAGCTGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCGCAGGTTTGTGTGTAAATTGCC 3559
Db  43142  --------TCAAGTCTGTATGTGAGTTCTGAGTGCTGAACTGGGGGTTCTGTGTGTGAAAATCTGC 43093

Qy   3560  CAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGATTGGATGAATGGCCTGCATCT 3619
Db  43092  CTGAGGCGGCAGGAGAATCACTGCCATTGCGTGAGCGAGGTTGGATGTTGGATGGCACTGGT 43033

Qy   3620  CAAGGACCATGGAAAATAGAAATGGAGACTCTATATGTGTCTCAAGGTAAGGTAG---CA 3676
Db  43032  C-AGGAGCATTAGGGAAGGAGGGGACTCCCAGACGTGTCCCCAAACCAGGGTGGGCCTCA 42974

Qy   3677  AGGTCTTTGGAGGGACACCTTGTCTAGAGATGTGGGCAACAGACTACAGACAGTATCTGT 3736
Db  42973  AGACCTTGGGAGGAGAACACTTGTCTGAAGAAGTTGGGGAACAG-----AAGGAGACCAGGCAT 42919

Qy   3737  ACAGAGCTAAGCAGCAGAGAGAGGGGGGTGTAGAATTCTCTTACTACTATCAAAAGGGAAAACTGAG 3796
```

```
Db    42918  GGCACTTATGCAGACTGAGGCCAGGACAGA------ATTTCCTGACAAAAGAAAACTGAG  42865
                     ||||| ||| |||||||||||  |||||| |||  |||||||||||||||||
Qy     3797  TCGTGCACCTGCAAAGTGGATGCGTCTCCCTAGACATCATGACTTTGTCTCGGGGAGCCA   3856
                     ||||| ||| |||||||||||||||||||||||||  |||  |||||||||||||||
Db    42864  CCATGGAGATGGCAACAGATCCCTTCCTGGCACCAT-ACTGCAGCTTTTAGTCTCCCTA   42806

Qy     3857  GCACTGTGGAACTTCAGGTCTGAGAGAGTAGGAGGCTCCCCTCAGCCTGAAGCTATGCAG   3916
                     |||||||| ||||||||||||||||||||||||||| |||||||||||||||||||
Db    42805  GCACTGGGGGCTCCAGTACTAACAGCAGGAAGAGATGCTCCCCAGCCTGGGGACTGTGTGAGGGG  42746

Qy     3917  ATAGCCAGGGTTGAAAGGGGAGAAGGGAGAGCCTGGGATGGGAGCTTGTGTTGTTGGAGGCA   3976
                     ||||| |||||||||||||||||||||||||||||||||||||||||
Db    42745  AGGTCAGAATGGGAGGGAAGGAGGAGGAGGCTGGGAACAGGGGAGGAAAGCCCATGGTTGGGAGGCG  42686

Qy     3977  GGGGACAGACAGATATTAAGCCTGGAAGAGAGGTGACCCTTACCCAGTTGTTCAACTCACCCT   4036
                     |||||| |||  ||| |||||||||||||||||||||||||  |||||||||||
Db    42685  GAGGACAGGCATTTGGCCTGCAGGAGAGGTGACCCTCACCCATGTTTTCAGTTCACCCT   42626

Qy     4037  TCAGATTAAAAATAACTGAGGTAAGGGC------CTGGGTAGGGAGGTGGTGTGAGACGC   4091
                     ||||||||||||||||||||||||||||||  |||||||||||||||||||||||||
Db    42625  TCGGGTTAAAAATAACTGAGGTAAGGGCCATGGCAGGGTAGGGGTGGGAGGCGGGTGTGAGAAGG  42566

Qy     4092  TCCTGTCTCCTCCCTGCATGCCCTGAGGCCCTTGGGAGGGAGGAGAAATGTGCCCAAGGACT   4151
                     |||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db    42565  TCCTGTCTCTTCCCACTATCTGCTCATCAGCCCTTTGAAGGGAGGAGAATGTGCCCAAGGACT  42506

Qy     4152  AAAAAAGGCCATGGAGCCAGAGGGCGCGAGGGCAACAGACCTTTCATGGGCAAACCTTGG   4211
                     ||||||||||| |||||||||||||||||||||||  ||| |||||||||||||||||||
Db    42505  AAAAAAGGCCGTGGAGCCAGGAGCCAGCAGAGAGGCTGGGCGCAGACCTTTCTTGGGCAAATCAGGG  42446

Qy     4212  GGCCC   4216
             |||||
Db    42445  GGCCC   42441
```

RESULT 14

```
CNS0000B/c
LOCUS       CNS0000B    196292 bp   DNA   linear   PRI   22-MAY-2001
DEFINITION  Human chromosome 14 DNA sequence BAC R-124D2 of library RPCI-11
            from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION   AL049829
VERSION     AL049829.4   GI:8217859
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196292)
  AUTHORS   Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Eckenberg,R., Bruls,T., deberardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
  TITLE     Sequencing of the human chromosome 14
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196292)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAY-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     On Jun 3, 2000 this sequence version replaced gi:6138746.
            ------------- Genome Center
            Center: Genoscope / Centre National de Sequencage
            Center code: GS
            Web site: http://www.genoscope.cns.fr/
            Contact: SeqRef@genoscope.cns.fr
            -------------
            The following BAC sequence is oriented from the T7 to the SP6 end.
            Upstream BAC (overlapping the T7 end) : R-244E17
            Downstream BAC (overlapping the SP6 end) : C-2201G16 (AC=AL132855)
            ------------- Summary Statistics
            Assembly program: Phrap; version 2.0
            Quality coverage: 7.94x in Q20 bases; sum-of-contigs
```

```
------------------------
Overall quality chart :
Range    :    bases
------------------------
  0 -  9 :
  1 -  9 :
 10 - 19 :
 20 - 29 :             10
 30 - 39 :             30
 40 - 49 :            829
 50 - 59 :           1909
 60 - 69 :           5193
 70 - 79 :          18093
 80 - 89 :          58972
 90 - 99 :         111256

Percentage of bases with a quality value >= 40 : 99 %.
```

```
FEATURES             Location/Qualifiers
     source          1..196292
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="R-124D2"
                     /clone_lib="RPCI-11"
     STS             78259..78398
                     /note="matching EMBL:AA908790
                     RHdb:RH102162
                     dbSTS:STS69699"
     STS             80433..80566
                     /note="matching EMBL:AA167748
                     RHdb:RH98727
                     dbSTS:STS68485
     STS             82164..82299
                     /note="matching EMBL:H79035
                     RHdb:RH78773
                     dbSTS:STS55514
     STS             93580..93808
                     /note="matching EMBL:AA007328
                     RHdb:RH75103
                     dbSTS:STS52191
     STS             107758..107951
                     /note="matching EMBL:R94929
                     RHdb:RH65111
                     dbSTS:STS45044
     STS             107996..108181
                     /note="matching EMBL:D11677
                     RHdb:RH47139
                     dbSTS:STS40201
     STS             128387..128525
                     /note="matching EMBL:H22648
                     RHdb:RH53520
                     dbSTS:STS25817
     STS             128419..128543
                     /note="matching EMBL:R87257
                     RHdb:RH53698
                     dbSTS:STS18321
     STS             138462..138738
                     /note="matching EMBL:R59134
                     RHdb:RH53972
                     dbSTS:STS42930
     STS             138509..138658
                     /note="matching EMBL:M78864
                     RHdb:RH95543
```

```
              RHdb:RH1243
              dbSTS:STS24710
              Identified using the e-PCR software (G. Schuler)"
STS           138568..138677
              dbSTS:STS20163
              Identified using the e-PCR software (G. Schuler)"
              138568..138647
              dbSTS:STS20162
STS           138805..138926
              /note="matching EMBL:H72023
              RHdb:RH68657
              dbSTS:STS48564
STS           161617..161761
              /note="matching EMBL:AA452257
              RHdb:RH92419
              dbSTS:STS65422
              Identified using the e-PCR software (G. Schuler)"
              194804..194903
              dbSTS:STS3407
STS           RHdb:RH76320
              Identified using the e-PCR software (G. Schuler)"

ORIGIN

Query Match        9.7%;    Score 558.2;  DB 9;  Length 196292;
Best Local Similarity  63.7%;   Pred. No. 3.6e-164;
Matches 1380;  Conservative  0;  Mismatches 658;  Indels 127;  Gaps  30;
```

```
Qy   2096 CTCACCTGTGGCCAGAGGGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGC 2155
          ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db 192568 CTCACCTGTGGCCGAAGCCTCCATCCCTGTCTTGTCCTTTATATTCCAGACTCCCAGAATGA 192509

Qy   2156 AGGGCT-------GGCTTGGGAGGCAGCTGGAAAGAGTATGTGAGAGCCAGGGGACA 2207
          |||||       |||||||||||||||||||||||||||||| | |||||||||||
Db 192508 GGAGCTGTGTGCAAGGCAGCTTAGGGAGGACCGGAGAGCAATATGTGAGAGCAGGACAGGGAGA 192449

Qy   2208 AGGGGGCCTAGGAAGAAGAGGAGAGGGCAAACCAGGCCACAAGAGGGCAGCCCAGA 2267
          ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db 192448 A-GGAGCTGGGGAAGGGAGGAGGGAGGAGGAGAGGGGTACAGCGGCAGAC---GGGCAGGGCCTGGG 192394

Qy   2268 ACTGAGTTAAACTCCTTCCTTGTTGCATCTTCATAGGAGGCAGTGGGAACTCTGTGACCA 2327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 192393 GCTGAGTGGAACACCTTCCTTCCTTCCTTCCTTCCACAGGAGGCGGGGGGGGACTACCCCCCACCAC 192334

Qy   2328 CCATCCCCCATG-------AGCCCCCACTACCCAAGTTTGGCCTGAGTGGCCATTCT 2381
          |||||||||||       || |||||||||||||||||||||||||||||||||||
Db 192333 CTCCACTCCCCCGGCTCGGCGAATTCTGGGGCCTCGAGTCCCTGCTCAGAGCACA 192274

Qy   2382 AGGTTCCCTGAGGACAGAGCCTGGCCTTTGTCTTGTTGTCACCTGACCCAGACTGACCTGGC 2441
          |||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||||
Db 192273 GGGTTCCCTGAGGACAAAGCCTGGCTGCTTCTTGTTGTCACCTGACCCAGACTGACCTGGC 192214

Qy   2442 GTTCTCAGTACCTTTGAATGCCCTCAAGAGCTTGAGAACCAGGCAGTGACATATTAGGCC 2501
          ||||||||||||||||||| |||| ||||||||||||||||| | ||||| |||||||||
Db 192213 GTTCTCAGCTCCTTATCATGTCCCAGAGCTATGGGAACCAAGCAGTGACGCACT-GGCC 192155

Qy   2502 ATGGGGCTAACCCTGGAGCTTGCACACAGGAGGCTCAGTGACCTCAGGGGACACAGCTGC 2561
          ||||||||||||||||| |||||||||||||||||||||||| ||  |||||||||||||
Db 192154 ATGGGGCTAACCCCCGGCTCGAATTCTGGGGGCCTCGAGTCCCTCAGTAACCT-------AGAGCTGG 192103

Qy   2562 AGACAGGTGGCCTTTATCCCCAAAGAGCAACCATTTGGCATAGGTGGCTGC---AAATGG 2618
          ||||||||||||||| ||||||||||||||||| |||||||| ||| || |   |||||
Db 192102 GGACAGGTGGCCTTGCATCCCCGTGGGGGGGTAATGATTTGGCATAGCACCTGCGCAGAGATGG 192043

Qy   2619 GAATGCAAGGTTGAATCAGGTCCCTTCAAGAATACTGCATGCAAGACCTAAGACCCCTGG 2678
          ||||||||||| || ||| ||||  |||||||| ||| ||   |||| ||| ||  |||
Db 192042 GGATGTGGGGGTCAACTAAGGTCCCCTCAAGGACCATTG-AGAAAAGGTCCAAATGCTCTGG 191984
```

```
Qy   2679 AGAGAGGGGTATGCTTCTCGCCCCACCCACCATAGGGGGAGTGAACTATCCTAGGGGGCT 2738
          |||||||||||||||||||||||||||||||||||||       ||||||||||||||||
Db 191983 AAAGAGAGGTGTGTTCCTAGCCCCA------GGAATGGGGTGGCAGCGTCCTAGGGAGCT 191930

Qy   2739 GGGCACCTGGGAGACACACATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCC 2798
          ||||||||||||||||||  ||||||||||||||||||||||||||||||||||| -
Db 191929 AGGG-----GAGAGCCACATCTTAATGCAAGTGCTCCTACCCAGAGAAGCTGACCCCT 191875

Qy   2799 TGTGTCCTGCCCACCTCCACACTCTAGAGCTATTGAGAGGTGACAGTAGATAGGGTGG 2858
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 191874 GTCGTCCTCCACCCACCTCCACACCCTGGAGCTATATTGAGAGGTGACAGTAAACTGGTGGG 191815

Qy   2859 GAGCTGGTAGCAGGGAGAGTGTTCCTGGGTGTGAGGGGTTAGGGGAAAGCCAGAGCAGGG 2918
          || |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db 191814 GG-GGGGGGCTGGGAGGGGTGTTCCTGGGTGTGAGGGGTG-GGGGGAAGGCCAGAGCAGGG 191759

Qy   2919 GAGTCTGGCTTGTCCTGAACACAATGTCTACTAGTTATAACAGGCATGACCTGCTA 2978
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 191758 GAGTCTGGCTTGTTCCTGAACACAATGTCCACTAGTCATAACAGGAACAGCCTGCTG 191699

Qy   2979 AAGACCCAACATCTACGACCTTGAAAAGACAGCAGCCCTGGAGGA------CAGGGGGTGT 3034
          |||||||||| ||| | ||||||||||||||||||||||||||||      |||||||||
Db 191698 AAGACCTGACCACCTACAGCGCCTGTG-AGAGATGGGACGCATTGGACGATGGGGAGCGCGGGGCC 191640

Qy   3035 CCTGAGCCTTGGTGCTTGATGTGTCCACAAAGGAGGGCATGAGTGTGAGTATAAGGCC 3094
          |||||||||||| | ||| |||||||||| ||||| |||||||||||||| ||||| ||
Db 191639 TTGAGGGGCCTTAGGGTGCTTGAGGACACCCCACAAAGAAGG---GAATGTGAGTATAAGCCC 191583

Qy   3095 CCAGGAGC---------GTTAGAGAAGGGCACTTGGAAGGGGTCAGTCTGCAGAGCCC 3144
          ||||||||         |||| ||||| || |||||||||||| | ||||||||||||
Db 191582 TGGAGGGGCTTGGAAAGGGCAACAAAATTCTGCCTGGGGAGGGGGCAGGTGCAGGAGCCC 191523

Qy   3145 CTATCC-ATGGAATCTGGAGCCTGGG--CCAACTGGTGTAAATCTCTGGGCCTGCCAGGC 3202
          | |||| |  ||||||| | ||| ||  ||||||| ||||| ||||||||||||||||| -
Db 191522 CACCCCGATGGAATGCAGTGCCTGGGGACCCAGGATGGTGTAAATCCCAGGCCCTGCCAG-- 191465

Qy   3203 ATTCAAAGCAGCACCTGCATCCTCTGGCAGCTGGCCCCTCCGGAGGCGCAACCCCCA 3262
          ||||| |||| | |||| ||||||||| |||||||||||||||||||||||||||| |
Db 191464 --------GCACCCTGCGGCTGCAGCCCCGGTGAGGCGGAGGCAGCAGAGGGGACAGCCCCTCA 191415

Qy   3263 CTTATACCCTTTCTCCCCTCAGCCCCAGGATTAACACCCTGGCCTCAGGCCTCACCT 3322
          | ||| |||| ||| |||||||||||  ||||||||||||||||||||||||||||||
Db 191414 CCACCCCCTTCCTCTCTCTCAGGCCCTAGCATTAGCACCATCTTCTT----CTGCCCACCC 191359

Qy   3323 CCCATCAGGAGTGGAGGGTTGCAGAGGGAGGGTAAAAACCTACATGTCCAAACATCATGG 3382
          |||||||||||||| |||| ||||| ||||| |  ||||  ||| || |||||||||||
Db 191358 CACCCCCAGGAGCAGCAGAGGGTAGCAG-GGGAGGCGGCAAAAGCCTATGTGTGCAAATGCGGATG 191300

Qy   3383 TGCACGATATATGGATCAGTATGTGTAGAGCA---AGAAAGGAAAATCTGCAGGCTTAAC 3439
          ||||  | ||| || ||||  || || ||| ||   ||||||| | |||| ||| |||| -
Db 191299 TGCA-AATGTGCATCCATTGACCAAAAAGAAAGTAGGAAGAGAAATCTCCAAGCTTGAG 191241

Qy   3440 TGGGTTAATGTGTAAAGTCTGTGTCCATGTGGTTAATGTGCATCTGAAAACGGGCATGGCT 3499
          ||||| || ||||||||| | |||||||||| ||| |||||| |||||| |||||||||||
Db 191240 TGGGTAAGTGTGAAAAAATCTGCATGTGTGGCTGAAGATGGCACAGACACG------- 191189

Qy   3500 GTGCAGCTCGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCAGGTTTGTGTGTAAATTGCC 3559
          ||||||||||||||||||||| || | |||| ||||||||||||||||||||||| |||||
Db 191188 --------TCAAGTCTGTATGTGAGACTGCTGAACTGGGGTTCTGTGTGTCGAAAATCTGC 191139

Qy   3560 CAAGGCAAAGTGGTGAATCCCTTCCATGGTTTAAAGAGATTGATGGATTGGCCTGCATCT 3619
          |||||||| | ||| |||||  |||||| || ||||| || | ||||  ||||||  |||
Db 191138 CTGAGGCGGCAGGGAGAAATCACTGCCATTGCGTGGAGCAGTTGGATGTTGGCCACTCTAT 191079

Qy   3620 CAAGGACCATGGCGAAAATAGAATGGCATTGGCATAGGTGGCTGC-----------CA 3676
          || ||| | | | |  | ||||||||  ||||||||| ||||  |           ||
Db 191078 C-AGGAGCATTAGGGAAGGGGTGGGACTCCAGACGTGTCCCCAAACCAGGGTGCCTCA 191020

Qy   3677 AGTGTCTTTGGGATCAGGTCCCTTCAAGAATACTGCATGCAAGACCTACAGTATCTGT 3736
          |||||||||||||| | | |||| ||||||| |  |||||| |||| |||  |  |||||
Db 191019 AGACCTTGGGAGAACACTTGCTGTGAAGACTTGGGGAACAG-----AAGGAGACCAGGCAT 190965
```

Qy   3737 ACAGAGTAAGGAGGAGAGGAGGGGTGTAGAATTCTCTTACTATCAAAGGGAAACTGAG 3796
              |||||| |||||| |||||||| || |||||||||||||||||
Db   190964 GGCACTTATGCAGACTGAGGCCAGGACAGA------ATTTCCTGACAAAGAGAAACTGAG 190911

Qy   3797 TCGTGCACCTGCAAAGTGGATGCTCTCCTAGACATCATGACTTTGTCTCTGGGGAGCCA 3856
              | |||| ||||| ||||||||||||||||||| || ||| | ||
Db   190910 CCATGGAGATGGACAACAGATCCCTTCCCTGGGCACCAT-ACTGCAGCTTTTAGTCCCTA 190852

Qy   3857 GCACTGTGGAACTTCAGGTCTGAGAGAGTAGGAGGCTCCCCTCAGCCTGAAGCTATGCAG 3916
              |||||| || ||||| ||||| ||||||||| || |||||||||||||
Db   190851 GCACTGGGGGCTCCCAGTACTAACAGCAGGAAGATGCTCCCAGCCTGGGACTGTGTGAGGG 190792

Qy   3917 ATAGCCAGGGTTGAAAGGGGGAAGGGAGAGAGCCTGGGATGGGGAGCTTGTGTGTTGGAGGCA 3976
              |||||| ||||||||||| |||||||||||| ||||||||||||||||
Db   190791 AGGTCAGAATGGGAAGGGAAGGAGGCTGGGGAACACAGGGGAACAGCCCATGGTTCGGGAGGCG 190732

Qy   3977 GGGGAGCAGATATTAAGCCTGGAAGAGAGAGGTGACCCTTACCCAGTTGGTTCAACTCACCCT 4036
              | |||||| ||||||| |||||| ||||||||||||||||||||||||||||||
Db   190731 GAGGACAGGCATTTGGGCCTGCAGGAGGTGACCCTTCACCCCATGTTTTCAGTTCACCCT 190672

Qy   4037 TCAGATTAAAAATAACTGAGGTAAGGGC------CTGGGTAGGGGAGGGTGGTGTGAGACGGC 4091
              ||||||||||||||||||||| |||||||||| ||||||||||||||||
Db   190671 TCGGGTTAAAAATAACTGAGGTAAGGGCCCATGGAGGGTGGGAGGGTGGGAGAAGG 190612

Qy   4092 TCCTGTCTCTCCTCTGCGATGCCCTGAGGCCCTTTGGGGAGGAGGAATGCCCAAGGACT 4151
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db   190611 TCCTGTCTCTCCCACTATCTGCTCATCAGCCCTTTGAAGGGGAGGAATGTGCCCAAGGACT 190552

Qy   4152 AAAAAAAGGCCATGGAGCCAGAGGGGCGGAGGGCCAACAGACCTTCATGGGCAAAACCTTGG 4211
              |||||||||||||||||| |||||| |||||||||||||||||||||||||||||||
Db   190551 AAAAAAAGGCCGTGGAGCCAGAGGGCTGGAGGCTGGGCAGCAGACCTTTCTTGGGCAAATCAGGG 190492

Qy   4212 GGGCC 4216
              |||||
Db   190491 GGGCC 190487

FEATURES   Location/Qualifiers
     source   1..31462
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="Cos HM-1 and Cos HM-13"
                     /cell_type="leukocyte"
                     /clone_lib="Drs Y.W. Kan and L. Chan"
     mRNA   join(4485..4506,5152..5184,5522..5735,6987..7130,
                     7378..7534,7630..7657,7935..8046,8367..8459,8976..9039,
                     9308..9406,9954..10057,10150..10288,11771..12039,
                     12322..12492,13709..14018,15118,15188,15467..15554,

                                                                    15644..15761,15906..16029,16303..
                                                                    18815..19057,19205..19381,19666..
                                                                    22275..22664,23001..23127,23209..
                                                                    24380..24563,24791..24956,25061..
                                                                    26570..26773,27658..27783,27990..
                                                                    30071..30205,30569..30643)
     exon   /number=1
                     4485..4506
     exon   /number=2
                     5152..5184
     exon   /number=3
                     5522..5735
     CDS   join(5535..5735,6987..7130,7378..7534,7630..7657,
                     7935..8046,8367..8459,8976..9039,9308..9406,9954..10057,
                     10150..10288,11771..12039,12322..12492,13709..14018,
                     15118..15188,15467..15554,15644..15761,15906..16029,
                     16303..16439,18400..18815,19057..19205,19381..
                     19666..19811,20071..23327..23663,23859..24380..24956,
                     23209..23327,23663,23859..24380,24563,24791,24956,
                     25061..25185,26079,26387,26570,26773,27658,27783,
                     27990..28265,29313,29408,30205,30569..30592)
                     /codon_start=1
                     /product="cardiac alpha-myosin heavy chain"
                     /protein_id="CAA79675.1"
                     /db_xref="GI:297024"
                     /db_xref="GOA:P13533"
                     /db_xref="UniProt/Swiss-Prot:P13533"
                     /translation="MTDAQMADFGAAAQYLRKSEKERLEAQTRPFDIRTECFVPDDKE
                     EFVKAKILSREGGKVIAETENGKTVTVKEDQVLQQNPPKFDKIQDAMLTFLHEPAVL
                     FNLKERYAAWMIYTYSGLFCVTVNPYKWLPVYNAEVVAAYRGKKRSEAPPHIFSISDN
                     AYQYMLTDRENQSILITGESGAGKTVNTKRVIQYFASIAAIGDRGKKDNANANKGTLE
                     DQIIQANPALEAFGNAKTVRNDNSSRFGKFIRIHFGATGKLASADIETYLLEKSRVIF
                     QLKAERNYHIFYQILSNKKPELLDMLLVTNNPYDYAFVSQGEVSVASIDDSEELMATD
                     SAFDVLGFTSEEKAGVYKLTGAIMHYGNMKFKQKQREEQAEPDGTEDADKSAYLMGLN
                     SADLLKGLCHPRVKVGNEYVTKGQSVQQVYYSIGALAKAVYEKMFNWMVTRINATLET
                     KQPRQYFIGVLDIAGFEIFDFNSFEQLCINFTNEKLQQFNHHMFVLEQEEYKKEGIE
                     WTFIDFGMDLQACIDLIEKPMGIMSILEEECMPKATDMTFKAKLYDNHLGKSNNFQK
                     PRNIKGKQEAHFSLIHYAGTVDYNILGWLEKNKDPLNETVVALYQKSSLKLMATLFSS
                     YATADTGDSGKSKGGKKKGSSFQTVSALHRENLNKLMTNLRTTHPHFVRCIIPNERKA
                     PGVMDNPLVMHQLRCNGVLEGIRICRKGFPNRILYGDFRQRYRILNPAAIPEGQFIDS
                     RKGTEKLLSSLDIDHNQYKFGHTKVFFKAGLLGLLEEMRDERLSRITTRMQAQARGQL
                     MRIEFKIVERRDALLVIQWNIRAFMGVKNWPWMKLYFKIKPLLKSAETEKEMATMKE
                     EFGRIKETLEKSEARRKELEEKMVSLLQEKNDLQLQVQAEQDNLNDABERCDQLIKNK
                     IQLEAKVKEMNERLEDEEEMNAELTAKKRKLEDECSELKKDIDDLELTLAKVEKEKHA
                     TENKVKNLTEEMAGLDEIIAKLTKEKKALQEAHQQALDDLQVEEDKVNSLSKSKVKLE
                     QVDDLEGSLEQEKKVRMDLERAKRKLEGDLKLTQESIMDLENDKLQLEEKLKKKEFD
                     INQQNSKIEDEQALALQLQKKLKENQARIEELEEELEAERTARAKVEKLRSDLSRELE
                     EISERLEEAGGATSVQIEMNKKREAEFQKMRRDLEEATLQHEATAAALRKHADSVAE
                     LGEQIDNLQRVKQKLEKEKSEFKLELDDVTSNMEQIIKAKANLEKVSRTLEDQANEYR
                     VKLEEAQRSLNDFTTQRAKLQTENGELARQLEEKEALISQLTRGKLSYTQQMEDLKRQ
                     LEEEGKAKNALAHALQSARHDCDLLREQYEEETEAKAELQRVLSKANSEVAQWRTKYE
                     TDAIQRTEELEEAKKKLAQRLQDAEEAVEAVNAKCSSLEKTKHRLQNEIEDLMVDVER
                     SNAAAALDKKQRNFDKILAEWKQKYEESQSELESSQKEARSLSTELFKLKNAYEESL
                     EHLETFKRENKNIQEISDLTEQLGEGGKNVHELEKVRKQLEVEKLELQSALEAEAS
                     LEHEEGKILRAQLEFNQIKAEIERKLAEKDEEMEQAKRNHQRVVDSLQTSLDAETRSR
                     NEVLRVKKKMEGDLNEMEIQLSHANRMAAEAQKVQKSLQSLLKDTQIQLDDAVRANDD
                     LKENIAIVERRNNLLQAEELELRAVVEQTERSRKLAEQELIETSERVQLLHSQNTSLI
                     NQKKMKSDLTQLQSEVEEAVQECRNAEEKAKKAITDAAMMAEELKKEQDTSAHLERM
                     KKNMEQTIKDLQHRLDEAEQIALKGGKKQLQKLEARVRELEGELEAEQKRNAESVKGM
                     RKSERRIKELTVTEEDKKNLLRLQDLVDKLQLKVKAYKRQAEEAEQANTNLSKFRK
                     VOEHLDEAEERADIAESQVNKLRAKSRDIGAKQKMHDEE"
     exon   6987..7130
                     /number=4
     exon   7378..7534
                     /number=5
     exon   7630..7657
                     /number=6
     exon   7935..8046
                     /number=7
     exon   8367..8459
                     /number=8
     exon   8976..9039
                     /number=9

```
Qy  2216 TAGAAAG-GAAGAAGAGAGGCAAACCAGGCCACACAAGAGGGCAGAGCCCAGAACTGAGT 2274
              |||      |||||||||| |||||    ||||
Db  2375 AAGGAGCCCTGGGGAGGAGGGAGGAGCGGTACAGCCGCAGAGCGGGCAGGGCTGGGGCTGAGT 2434

Qy  2275 TAACTCCTTCCTTGTTGCATCTTCCTCCTTCCACAGGAGGCAGTGGAACTTCTGTGTGACCACCATCCC 2334
              ||||||||||||||||                        || |||   ||
Db  2435 GAACACCTTCCTTCCTCCTTCCTGTCCTTATATTCCACCCCACCACTTCCACTC 2494

Qy  2335 CCATGAGCCC------CCACTACCCATACCAAGTTGGCCTGAGTGGCATTCTAGGTTCCC 2389
              ||||               |||||| |||||||
Db  2495 CCCGCTGCCCTGCCACCACCATACCAAGTGAGGCCTGCTCAGAGCACAGAGGTTCCC 2554

Qy  2390 TGAGGACAGAGCCTTGGCCTTTGTCTCTTGACCTTGGACCTGACCCAAGCTGACCCAATGTTCTCAG 2449
              ||||||| |||||||||  ||  || || |||||||  |||| ||| ||| ||
Db  2555 TGAGGACAAAGCCTGGTCTTTGTCACCTGACCCCGACCAGCGACTGACCTGGCGTTCTCAG 2614

Qy  2450 TACCTTTGAAATGCCTCAAGAGCTTGAGAACCAGGCAGTGACATATTAGGCCATGGGCTA 2509
              ||||||||| || ||||||   ||||||||||| |||||||     |||||||||||||||
Db  2615 CTCCTTATCATGTCCCAGAGCTATGGGAACCAAGCAGTGACGCACT-GGCCATGGGCTA 2673

Qy  2510 ACCCTGGAGCTTGCACAGAGGAGCTCAAGTGACCTCCAGGGACACAGCTGCAGACAGGT 2569
              ||||                            |||  |||||||||||||
Db  2674 ACCCCCCCGCTCGAATTCTGGGGCCTGAGTAACCT------AGAGCTGGGGACAGGT 2725

Qy  2570 GGCCTTTATCCCCAAAGAGCAACCATTGGCATAGGTGGCTGC---AAATGGGAATGCAA 2626
              ||||  || ||||    |||| || |||| |||||| |||
Db  2726 GGCCTGCATCCCC--TGGGGTAATGATTTGGCATAAGCACCTGCAGAAGAGTGGGGATGTGG 2784

Qy  2627 GGTTGAATCAGTTCCTTCAAGATACTGCATGCAAGACCTTAAGACCTCAGAGAGAGGG 2686
              |||||       |||   ||| || |||| ||| ||||| ||||||||
Db  2785 GGTCAACTAAGGTCCCTGCCAAGGACATTG-AGAAAAGGTCCAAATGCTCTGGAAAGAGAG 2843

Qy  2687 GTATGTCTCTGCCCCCACCCACCCATAAGGGGAGTGAACTATCCTAGGGGGCTGGCGACCT 2746
              ||||| |||||            |||   ||  |||| |||| ||||||||
Db  2844 GTGTGTTCCTGAGCCCA-----GGAATGGGGTGGGCAGCGTCCTAGGGAGCTAGGG---- 2894

Qy  2747 TGGGGAGACACCACATTACTGAGAGTGCTGAGCCCAGAGAAAACTGACCGCCCTGTGTCCT 2806
              |||||||| ||| | |  |||||||| |||   ||||||||| ||||||||||  ||||||
Db  2895 -GAGAGGCCACATCTTAATGAGAAGTGCTCCTACCCAGAGAAGCGACCCCTGTCTGTCCT 2953

Qy  2807 GCCCACCTCCACACTCTAGAGCTATTATTGAGAGTGACAGTAGATAGGTGGAGCTGGT 2866
              ||||||||||| ||||| |||||| || || ||||||||||         || || ||||
Db  2954 ACCCACCTCCACACCCTGGAGCTATATTGAGAGTGACAGTAAACTG---GTGGGAGGGG 3010

Qy  2867 AGCAGGGAGAGTGTTCCTGGGTGTGTAGGGGAAAGCCAGACGAGCAGGGAGAGTCTGG 2926
              ||||||||||||| ||||| ||||  |||||   ||||||| ||||||| ||||||||
Db  3011 GGCTGGGAGGGTGTTCCTGGGTGTG-GGGGGTG-GGGGGAAGGGCCAGAGCAGGGGAGTCTGG 3069

Qy  2927 CTTTGTCTCCTGAACACAATGTCTACTTAGTTATAACAGCATGGAGATGGGGAGCCTAAAGACCCA 2986
              |||| ||||||||| || |||| ||||| || ||| |||| ||||| |||| || |||||
Db  3070 CTTTGTTTCCTGAACACAATGTCCACTTAGTCATAACAGGCAGGAGGATGGGAAGAACGCTGCTGAAGACCTG 3129

Qy  2987 ACATCTACCACCTGAAAGACAGCACCCTGAGGACAGAGGGTTGTGTTCCTGGAGCCCTTG 3046
              || ||||| |||| ||||| |||| |||||||||| |||| |||    ||||| ||||
Db  3130 ACACCTACAGCCTCTG-AGAGTGGCAGCATTGGAGGATTGGGAGGCGGGTATAAGGGCGGGGC 3188

Qy  3047 GGTGCTTGATGGTGCCACAAAGGAGGGCATGAGTGTGAGTAAGGGC------ 3102
              |||||||| |  ||||| |||||||| |  || || ||||||||||
Db  3189 TAGGGTGCTTGAGGAGACACCCCACAAAGAGGAATGTGAGTAATGTGAGTATAAGGGGCTTGG 3248

Qy  3103 ------GTTAGAAGGGCACTTGGAAGGGGGTCAGTCTGCAGAGCCCCTATCCATGGAA 3156
                    |||| || ||| |||||||||||    ||| ||||||||||
Db  3249 AAAGGGCAACAAAATTCTGCCTGGGAGGGGGTCGCAGCGCCCACCCGATGGAA 3308

Qy  3157 TCTGGAGCCTGGGG--CCAACTGGTGTAAATCCTGGGCCTGCCAGGCATTCAAAGCAGCA 3215
              ||||||||||||||                       ||||||||||
Db  3309 TGCAGTGCCTGGGGACCAGATGGTGTAAATCCCCAGGCCTGCCAG---------GCA 3356

Qy  3216 CCTTGCATCCTCTGCAGCCTGGGAGGCCGGAGGGGGAGGAGACAAGGGGGCC 3275
              || || ||| |||||||| |||| ||| ||||  || |||| |||
Db  3357 CCTGCACCCTCTGCACCCCGTGAGGCAGAGAGCCCGGAGGAGAAGGGAGAGCAGGGGGAGAG 3416
```

```
exon   9308..9406
       /number=10
exon   9954..10057
       /number=11
exon   10150..10288
       /number=12
exon   11771..12039
       /number=13
exon   12322..12492
       /number=14
exon   13709..14018
       /number=15
exon   15118..15188
       /number=16
exon   15467..15554
       /number=17
exon   15644..15761
       /number=18
exon   15906..16029
       /number=19
exon   16303..16439
       /number=20
exon   18400..18655
       /number=21
exon   18815..19057
       /number=22
exon   19205..19381
       /number=23
exon   19666..19811
       /number=24
exon   20071..20161
       /number=25
exon   22275..22664
       /number=26
exon   23001..23127
       /number=27
exon   23209..23327
       /number=28
exon   23663..23859
       /number=29
exon   24380..24563
       /number=30
exon   24791..24956
       /number=31
exon   25061..25185
       /number=32
exon   26079..26387
       /number=33
exon   26570..26773
       /number=34
exon   27658..27783
       /number=35
exon   27990..28265
       /number=36
exon   29313..29408
       /number=37
exon   30071..30205
       /number=38
exon   30569..30643
       /number=39
ORIGIN
```

```
Query Match     8.6%;  Score 494.8;  DB 9;  Length 31462;
Best Local Similarity  62.2%;  Pred. No. 3.2e-144;
Matches 1340;  Conservative  0;  Mismatches 702;  Indels 114;  Gaps  31;

Qy  2096 CTCACCTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGC 2155
         |||||||||||||   |||||||||  |||||     ||| | |||| |||||
Db  2256 CTCACCTGTGGCCGAAGCCTCCATCCCTGTCCTTATATTCCAGGACTCCCAGCAGAATGA 2315

Qy  2156 AGGGCTGCTTGGGAGGCAGCTGGAAAGAGTATGTGAGAGCCAGGACCAAGGGGGCC 2215
         | |||||||| ||||||||||  |   |||| ||| || |||| ||
Db  2316 GGAGCTGTG-TGCAAGGCAGCTTAGGGAGGCAGCGAGGACGGAGAGCGGGGGAG 2374
```

```
Qy   3276  TCCCTCAGCCCCAGGATTAACACCTCTGGCCTTCCCCCTTCCCACCTCCCATCAGGAGTG  3335
                 ||||||||||||    ||||  ||      |       |||        ||||||
Db   3417  CTCTCAGCCCCTAGGATTAGCACCATCTTCTT-----CTGCCCACCCCACCCCAGGAGCA  3471

Qy   3336  GAGGGTTGCAGAGGAGGGTAAAAACCTACATGTCCAAACATCATGGTGCACGATATATG   3395
                 |||    |||  ||||||| |||||  ||||    |||        |||||| |||
Db   3472  GAGGGTAGCAG-GGGAGGGCAAAAGCCTATGTGTGCAAATGCGGATGTGCA-AATGTGTG  3529

Qy   3396  GATCAGTATGTGTAGAGGCA---AGAAAGGAAATCTGCAGGCTTAACTGGGTTAATGTGT  3452
                 ||||     |    ||||   |||||||  |||||||    ||   || |   ||||
Db   3530  CATCCATGTACACAAAGAAAGTAGAAAAGAAATCTGCAAGCTTGAGTGGGTAAGTGTGA  3589

Qy   3453  AAAGTCTGTGCATGTGTGTCTGACTGAAAACGGGCATGGCGTGTGCAGCTGTTCA      3512
                 || | |||||||||||||   || |||||||   |||  ||| ||||||||||  |
Db   3590  AAAATCTGCATGTGTGGCTGAAGATGGGCACACAGACACCGGTCAG-----------   3633

Qy   3513  GTTCTGTGCGTGAGGTTACCAGACTGCAGGTTTGTGTGTAAATTGCCCAAGGCAAAGTGG  3572
                 ||||   | |||||||||||    ||||||||||||   ||| |   ||||| |||
Db   3634  --TCTGTATGTGAGAGTGCTGAACTGGGGTTCTGTGTGAAAAATCTGCCTGAGCGCGCAGG  3691

Qy   3573  GTGAATCCCTTCCATGGTTTAAAGAGATTGGATGATGCCTGCATCTCAAGGACCATGGA  3632
                 |||||   ||||   |||| ||   |||||||| |   ||    || |||||||||
Db   3692  GAGAATCACTGCCATTGCGTGAGCAGGTTGGCATGTTGGCCACTCTATC-AGGAGCATTAG  3750

Qy   3633  AAATAGAATGGACACTCTATATGTGTCTTCTAAGCTAAGGTAG---CAAGGTCTTTGGAGG  3689
                 || | ||||| ||||| |||||| || |   |||   || ||   |||||| |||
Db   3751  GGAAGGGGTGGGGACTCCAGACGTGTCCCCAAACCAGGTGCCTCCAAGACCTTGGGAGA  3810

Qy   3690  ACACCTGTCTAGAGATGTTGGGCAACACAGACTACAGACATCTGTACAGAGTAAGGAG  3749
                 ||| |||||||    |||||||    ||  |  ||| ||||||| || ||||||||
Db   3811  ACACTTGTCTGAAGACTTGGGAACAG-----AAGGAGACCAGGCATGCCACTTATGCAG  3865

Qy   3750  AGAGAGAGGGGGTGTAGAATTCTCTTACTATCAAAAGGGAAACTGAGTCGTGCACCTGCA  3809
                 | ||                 |||   ||| |||||  ||||||||  ||  |||||||
Db   3866  ACTGAGGGCCAGGACAGAA-----TTTCCTGACAAAAGAAAACTGAGCCATGGAGATGA  3919

Qy   3810  AAGTGGATGCTCTCCCTAGACATCATGACTTTGTCTCTGGGAGCCAGCACTGTGGAACT  3869
                 ||| |||||||   |||     ||| ||   |||   ||     ||| ||| ||| |||
Db   3920  CAACAGATCCCTTCCCTGGGCA--CATACTGCAGCTTTTAGTCCCTAGCACTTGGGGCTC  3977

Qy   3870  TCAGGTCTGAGAGTAGGAGGCTCCCCTCAGCCTGAGCTATGCAGATAGCCAGGGTTG  3929
                 ||||||||  |||||||||  |  |||   || ||  || ||  | |   ||||||||
Db   3978  CAGTACTAACAGCAGGAGATGCTCCCAGCCTTGGGACTGTGTGAGGGATGTCAGAATGG  4037

Qy   3930  AAAGG-GGGAAGGGAGAGCCTGGGATGGGAGCTTGTGTTGGAGGCAGGGGACAGATAT  3988
                 |||||  |||| || ||| ||||||| |||||  |||||| ||||| |  |||||||||
Db   4038  AAAGGAGGAGGCTGTGGGAACAGGGAGGGAAAGCCCATGGTTGGGAGGCCGAGGACAGGCAT  4097

Qy   3989  TAAGCCTGGAAGAGAAGTGACCCTTACCCAGTTGTTCAACTCACCCTTCAGATTAAAAA  4048
                 ||||||||||||   ||||| ||  ||||||||||||  |  |||||  |||| |||||
Db   4098  TTGGCCTCGAGAGAAGGTGACCCTCACCCATGTTTC-AGTCTCACTTCGGGGGGAAAAA  4156

Qy   4049  TAACTGAGGTAAGGGC-----CTGGGTAGGGAGGGTGGTGGAGACGCTCCTGTCTCTCCC  4103
                 ||||||||| ||||||     |  ||| || |||||||||||  ||| || ||  |||||
Db   4157  TAACTGAGGTAAGGGCCCATGCAGGGGTTGGGAGAGGCCGGTGTGAAGAGTCCTGTCTTCCC  4216

Qy   4104  TC-TGCATGCCCTGAGGCCCTTTGGGGAGGAGG--AATGTGCCAAGGACTAAAAAAAGG  4160
                 || | |  |||||||||||| ||  |  |||||  |||||||||| |||| |||||||||
Db   4217  ACATATCTGCTCATCAGCCCTTTGAAGGGAGTAATGCCCTAAGGAGTAAAAAAAAGG  4276

Qy   4161  CCATGAGCCACAGGGGGCGAGGGCAACAGACCTTTCATGGGCAAACCTTGGGGCCC   4216
                 |  ||  |||||||||||||||||||||||||||||||  ||||||||||||||||
Db   4277  CCGTGGAGCCAGAGAGGCTGGGCGCAGCAGACCTTTCTTGGGCAAATCAGGGGGGCCC   4332
```

Search completed: April 25, 2005, 09:57:18
Job time : 16397 secs

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

```
            GenCore version 5.1.6
     Copyright (c) 1993 - 2005 Compugen Ltd.

Run on:    April 25, 2005, 02:58:32 ; Search time 2127.75 Seconds
                                     (without alignments)
                       16363.543 Million cell updates/sec

Title:          US-10-613-728-1
Perfect score:  5735
Sequence:       1 ggatcctgcaaggtcacaca........ccataagagtttgagtcgac 5735
Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0
Searched:       563728 seqs, 3035525691 residues
Total number of hits satisfying chosen parameters:     11267456
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
Database :      Published Applications NA:*
    1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
    2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
    3: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
    4: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
    5: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
    6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
    7: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
    8: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
    9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
   10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
   11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
   12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
   13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
   14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
   15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
   16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
   17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
   18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
   19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
   20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
   21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
   22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	5735	100.0	5735	17	US-10-613-728-1	Sequence 1, Appli
2	5045	88.0	5443	13	US-10-024-066-5	Sequence 5, Appli
3	5045	88.0	5443	16	US-10-332-966-2	Sequence 2, Appli
4	5045	88.0	5443	15	US-10-944-375-15	Sequence 15, Appl
5	1279.8	22.3	1679	17	US-10-027-655-5	Sequence 5, Appl
6	407.6	7.1	5190	18	US-10-798-037-1	Sequence 1, Appli
7	311.8	5.4	520	9	US-09-874-389-7	Sequence 7, Appli
8	311.8	5.4	520	10	US-09-921-650-7	Sequence 7, Appli
9	311.8	5.4	520	11	US-09-241-347-7	Sequence 7, Appli
10	311.4	5.4	990	15	US-10-134-643-17	Sequence 17, Appl
c 11	311	5.4	4438	16	US-10-102-143-12	Sequence 12, Appl
c 12	311	5.4	4479	16	US-10-102-143-10	Sequence 10, Appl
c 13	311	5.4	4556	16	US-10-102-143-8	Sequence 8, Appli
c 14	311	5.4	6346	16	US-10-102-143-6	Sequence 6, Appli
c 15	311	5.4	6423	16	US-10-102-143-4	Sequence 4, Appli
c 16	311	5.4	8287	16	US-10-102-143-14	Sequence 14, Appl
c 17	311	5.4	8364	16	US-10-102-143-16	Sequence 16, Appl
c 18	310.4	5.4	450	9	US-09-281-674-6	Sequence 6, Appl
c 19	310.4	5.4	450	9	US-09-777-317-6	Sequence 6, Appli
c 20	310.4	5.4	450	9	US-09-892-227-6	Sequence 6, Appli
c 21	310.4	5.4	450	9	US-09-874-389-9	Sequence 9, Appli
c 22	310.4	5.4	450	10	US-09-921-650-9	Sequence 9, Appli
c 23	310.4	5.4	450	11	US-09-241-347-9	Sequence 9, Appli
24	309.4	5.4	447	19	US-10-169-050-3	Sequence 3, Appli
25	309.4	5.4	450	9	US-09-281-674-5	Sequence 5, Appli
26	309.4	5.4	450	9	US-09-777-317-5	Sequence 5, Appli
27	309.4	5.4	450	9	US-09-892-227-5	Sequence 5, Appli
28	309.4	5.4	450	9	US-09-874-389-8	Sequence 8, Appli
29	309.4	5.4	450	10	US-09-921-650-8	Sequence 8, Appli
30	309.4	5.4	450	11	US-09-241-347-8	Sequence 8, Appli
31	309.4	5.4	456	9	US-09-900-530A-28	Sequence 28, Appl
32	309.4	5.4	470	16	US-10-375-884-2	Sequence 2, Appli
33	309.4	5.4	500	17	US-10-148-521-21	Sequence 21, Appl
34	309.4	5.4	2061	18	US-10-789-303-1	Sequence 1, Appli
35	309.4	5.4	2792	16	US-10-375-884-8	Sequence 8, Appli
36	309.4	5.4	3621	10	US-09-957-458B-7	Sequence 7, Appli
37	309.4	5.4	3752	10	US-09-957-458B-8	Sequence 8, Appl
38	309.4	5.4	3871	19	US-10-169-050-51	Sequence 51, Appl
39	309.4	5.4	3926	19	US-10-169-050-59	Sequence 59, Appl
40	309.4	5.4	3930	19	US-10-169-050-36	Sequence 36, Appl
41	309.4	5.4	4224	10	US-09-957-458B-10	Sequence 10, Appl
42	309.4	5.4	4382	19	US-09-957-458B-9	Sequence 57, Appl
43	309.4	5.4	4386	19	US-10-169-050-57	Sequence 57, Appl
44	309.4	5.4	4824	19	US-10-169-050-52	Sequence 52, Appl
45	309.4	5.4	4963	9	US-09-281-674-9	Sequence 9, Appli

ALIGNMENTS

```
RESULT 1
US-10-613-728-1
; Sequence 1, Application US/10613728
; Publication No. US2004010813A1
GENERAL INFORMATION:
  APPLICANT: Cincinnati Children's Hospital Medical Center
  APPLICANT: Robbins, Jeffrey
  TITLE OF INVENTION: A ROBUST, INDUCIBLE CARDIAC PREFERRED
  TITLE OF INVENTION: EXPRESSION SYSTEM FOR TRANSGENESIS
  FILE REFERENCE: CHM02 GN053
  CURRENT APPLICATION NUMBER: US/10/613,728
  CURRENT FILING DATE: 2003-07-03
  PRIOR APPLICATION NUMBER: 60/393,525
  PRIOR FILING DATE: 2002-07-03
  NUMBER OF SEQ ID NOS: 10
  SOFTWARE: FastSEQ for Windows Version 4.0
  SEQ ID NO 1
  ; LENGTH: 5735
  ; TYPE: DNA
  ; ORGANISM: Artificial Sequence
  ; FEATURE:
  ; OTHER INFORMATION: Inducible Cardiac preferred promoter
US-10-613-728-1

Query Match        100.0%;   Score 5735;   DB 17;   Length 5735;
Best Local Similarity  100.0%;   Pred. No. 0;
Matches 5735;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy    1 GGATCCTGCAAGGTCACACAAGGGTCTCACCCACCAGGTGCCCTAGTCTCAATTCAGT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GGATCCTGCAAGGTCACACAAGGGTCTCACCCACCAGGTGCCCTAGTCTCAATTCAGT 60

Qy   61 TTCCATGCCTTGTTCTCACAAATGCTGGCCTCCCCAGAGCTAATTGGACTTTGTTTTAT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy   1201 ACACTCGTGTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACCACTGAGCTAG 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACACTCGTGTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACCACTGAGCTAG 1260

Qy   1261 GCAACCAGAGCATGGGCCCGTGCTGAGGATGAGGAGAGTTGGTTACCAATAGCAAAAACAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCAACCAGAGCATGGGCCCGTGCTGAGGATGAGGAGAGTTGGTTACCAATAGCAAAAACAG 1320

Qy   1321 CAGGGGAGGGAGGAACAGAGAACAGAGAGGAAATAAGGAAGGAAGGAAAGGCCAGTCAATCAGA 1380
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 CAGGGGAGGGAGGAACAGAGAACAGAGAGGAAATAAGGAAGGAAGGAAAGGCCAGTCAATCAGA 1380

Qy   1381 TGCAGTCAGAAGAGATGGGAAGCCAACACACAGCTTGAGCAGAGAAACAGAAAGGGAG 1440
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 TGCAGTCAGAAGAGATGGGAAGCCAACACACAGCTTGAGCAGAGAAACAGAAAGGGAG 1440

Qy   1441 AGAATTCTGGGCATAAGGAGGCCACAGAAGAAGAGAGCCCCAGGCCCCCAAGTCTCCTCTTT 1500
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AGAATTCTGGGCATAAGGAGGCCACAGAAGAAGAGAGCCCCAGGCCCCCAAGTCTCCTCTTT 1500

Qy   1501 ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC 1560

Qy   1561 AGGAAGAGACCCGTAGGGAGGGGATCACTGGATGAAGAGATGTGTGGAGAAGTCCAGG 1620
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 AGGAAGAGACCCGTAGGGAGGGGATCACTGGATGAAGAGATGTGTGGAGAAGTCCAGG 1620

Qy   1621 GCAACCTAAGAGCCAGAGCCTGGAAACGATACCATAAAGAGCAAGATAAAGGTGGCCAGG 1680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCAACCTAAGAGCCAGAGCCTGGAAACGATACCATAAAGAGCAAGATAAAGGTGGCCAGG 1680

Qy   1681 CTGTGCACACAGAGGGTCGAGGACTGGTGTGTAGAGCCTCAAGATAAGGATGATGCTCAGA 1740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CTGTGCACACAGAGGGTCGAGGACTGGTGTGTAGAGCCTCAAGATAAGGATGATGCTCAGA 1740

Qy   1741 ATGGGCGGGGGGGGGGGGATTCTGGGGGGGGGGAGGAAGTGAGGAAGAACAG 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATGGGCGGGGGGGGGGGGATTCTGGGGGGGGGGAGGAAGTGAGGAAGAACAG 1800

Qy   1801 AGAATCTGAAGCCTGGAAACGATGGAAAGCTGGAAACAGGAACACCAGGCTACCTTTAGATG 1860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 AGAATCTGAAGCCTGGAAACGATGGAAAGCTGGAAACAGGAACACCAGGCTACCTTTAGATG 1860

Qy   1861 TAAATCATGAAAGACAGGAGAAGGAAGCTGGAGAGAGTAACAAATGGAAGAATGGTCCTTCTCAC 1920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 TAAATCATGAAAGACAGGAGAAGGAAGCTGGAGAGAGTAACAAATGGAAGAATGGTCCTTCTCAC 1920

Qy   1921 CATGGAAGCAAGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGAAGGCAGAG 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CATGGAAGCAAGACAAGCCAGGTTGAGCGCTCCGTGAAATCAGCCTGCTGAAGGCAGAG 1980

Qy   1981 CCCTGGTATGAGCAGCGGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGC 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 CCCTGGTATGAGCAGCGGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGC 2040

Qy   2041 GTAGACACAGGAACAGAGAGCGGGGACGCGGGGAGCCAGGTAACAAAGAAATGGTCCTTCTCAC 2100
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GTAGACACAGGAACAGAGAGCGGGGACGCGGGGAGCCAGGTAACAAAGAAATGGTCCTTCTCAC 2100

Qy   2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGGC 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 CTGTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAGGGGC 2160

Qy   2161 TGGCTTGGGAGGGAGCTGGAAAGTATGTGAGAGTCGAGAGACAAGGGGACCTAGGA 2220
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TGGCTTGGGAGGGAGCTGGAAAGTATGTGAGAGTCGAGAGACAAGGGGACCTAGGA 2220

Qy   2221 AAGGAAGAGAGGGCAAACCAGGCCACAAAGAGGGCAGAGCCCAGAACTGAGTTAACTC 2280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AAGGAAGAGAGGGCAAACCAGGCCACAAAGAGGGCAGAGCCCAGAACTGAGTTAACTC 2280
```

```
Db     61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTTGGACTTTGTTTTTAT 120

Qy    121 TTCAAAAGGGCCTGAATGAGGAGATCTTGTGCTACCAGCTCTAAGGGTGCCCGTGA 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTCAAAAGGGCCTGAATGAGGAGATCTTGTGCTACCAGCTCTAAGGGTGCCCGTGA 180

Qy    181 AGCCCTCAGACTTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCCCTCAGACTTGGAGCCTTTGCAACAGCCCTTTAGGTGGAAGCAGAATAAAGCAATTT 240

Qy    241 TCCTTAAAGCCAAAATCTGCCTCTAGAACTCTGCTTCTTCTGACCTCGGTCCCTGGGCTCT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCTTAAAGCCAAAATCTGCCTCTAGAACTCTGCTTCTTCTGACCTCGGTCCCTGGGCTCT 300

Qy    301 AGGGTGGGGAGGGTGGGGCTTGGGAAGAAGAAGGTGGGAAGTGGGCAAAAGCCGATCCCTAG 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGGGTGGGGAGGGTGGGGCTTGGGAAGAAGAAGGTGGGAAGTGGGCAAAAGCCGATCCCTAG 360

Qy    361 GGCCCTGTGAAGTTCGGAGCCTTCGGAGCCTTCCCTGTACAGCACTGCTCATAGATCCTCCTCCAGCC 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCCCTGTGAAGTTCGGAGCCTTCGGAGCCTTCCCTGTACAGCACTGCTCATAGATCCTCCTCCAGCC 420

Qy    421 AAACATAGCAAGAAGTGATACCTCCTTTGTGACTTCCCCAGGCCCAGTACCTGTCAGGTT 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAACATAGCAAGAAGTGATACCTCCTTTGTGACTTCCCCAGGCCCAGTACCTGTCAGGTT 480

Qy    481 GAAACAGGATTTAGAGAAGAGCCCTGAACTCACCTGAACTCTGAAGCTCACCACCAGGCA 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAAACAGGATTTAGAGAAGAGCCCTGAACTCACCTGAACTCTGAAGCTCACCACCAGGCA 540

Qy    541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAGCTATAATGCGGCCAC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAGCTATAATGCGGCCAC 600

Qy    601 AGAAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGGTATGACCCCCC 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGAAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTCAGTCGGCAAAGGTATGACCCCCC 660

Qy    661 TTCAGCAGATGTAGTAATGTCACCCCCTTAGATCCCATCCCAGGCAGGTCTCTAAGAGGACATG 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCAGCAGATGTAGTAATGTCACCCCCTTAGATCCCATCCCAGGCAGGTCTCTAAGAGGACATG 720

Qy    721 GGATGAGAGATGTAGTCAGCTGCTGCATTCCAAACACAGCTATCCACAGTGTCCCTGCCCC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGATGAGAGATGTAGTCAGCTGCTGCATTCCAAACACAGCTATCCACAGTGTCCCTGCCCC 780

Qy    781 TTCCACTTAGCCAGGACAGTAACCTTAGCCTATCTTCCTTCCTTCTTCCTCCATCCTCCCAG 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTCCACTTAGCCAGGACAGTAACCTTAGCCTATCTTCCTTCCTTCTTCCTCCATCCTCCCAG 840

Qy    841 GACCACCCCCTGTCTGCAGTATTCATTTCATTTCTTCACGTCGTGTCTGACTTCCAT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 GACCACCCCCTGTCTGCAGTATTCATTTCATTTCTTCACGTCGTGTCTGACTTCCAT 900

Qy    901 TTGCAAGGCTTTTGACCTTCTGCGACTGCTGGAAGATAGAGTTGGGCCCGTAGGTGTGGCAA 960
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTGCAAGGCTTTTGACCTTCTGCGACTGCTGGAAGATAGAGTTGGGCCCGTAGGTGTGGCAA 960

Qy    961 GCCATCTCAAGAGAAGCACAGGGGGACCAGGGGGACCAGATTTTTGGAAGATCAGGAACTAAA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCCATCTCAAGAGAAGCACAGGGGGACCAGGGGGACCAGATTTTTGGAAGATCAGGAACTAAA 1020

Qy   1021 TCACTGGCGGGGGGCCTCGGGGGGTAGAAAAAAGAGTGAGTGGTCCGCTCCAGCTAAGCCAAGC 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TCACTGGCGGGGGGCCTCGGGGGGTAGAAAAAAGAGTGAGTGGTCCGCTCCAGCTAAGCCAAGC 1080

Qy   1081 TAGTCCCCGAGATACTCTGCCACAGCTGGCTGGGGTAGCTTTTAGGAATGTGGGTC 1140
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TAGTCCCCGAGATACTCTGCCACAGCTGGCTGGGGTAGCTTTTAGGAATGTGGGTC 1140

Qy   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACTTCAC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACTTCAC 1200
```

Qy 3361 CCTACATGTCCAAACATCATGGTGCAGATATGGATCAGTATGTGTAGAGGCAAGAAA 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 CCTACATGTCCAAACATCATGGTGCAGATATGGATCAGTATGTGTAGAGGCAAGAAA 3420

Qy 3421 GGAAATCTGCAGGCTTAACTGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTGTCTG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GGAAATCTGCAGGCTTAACTGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTGTCTG 3480

Qy 3481 ACTGAAAACGGGCATGGCCATGGCGTGTGCAGCTGTTCAGTTCGTGTGCGTGAGTGAAT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 ACTGAAAACGGGCATGGCCATGGCGTGTGCAGCTGTTCAGTTCGTGTGCGTGAGTGAAT 3540

Qy 3541 GGTTTGTGTGTAAATTGCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GGTTTGTGTGTAAATTGCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600

Qy 3601 TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAATGCAGACACTCTATATGTGTCT 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAATGCAGACACTCTATATGTGTCT 3660

Qy 3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC 3720

Qy 3721 TACAGACAGTATCTGTACAGAGTAAGGAGGAGGAGGGGTGTAGAATTCTCTTACTA 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 TACAGACAGTATCTGTACAGAGTAAGGAGGAGGAGGGGTGTAGAATTCTCTTACTA 3780

Qy 3781 TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCTAGACATCATGACTT 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGCTCTCCCTAGACATCATGACTT 3840

Qy 3841 TGTCTCTGGGAGCCAGCACTGTGACGTCTCAGGTCTGAGAGAGTAGGAGGCTCCCCTCA 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTCTCTGGGAGCCAGCACTGTGACGTCTCAGGTCTGAGAGAGTAGGAGGCTCCCCTCA 3900

Qy 3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAGGGAGGGAGCCTGGGATGGGAGC 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAGGGAGGGAGCCTGGGATGGGAGC 3960

Qy 3961 TTGTGTGTTGGAGGCAGGGGACAGATATTAAGCCTGGAGGTAAGGGCCTGGGTAGGGAGGTG 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 TTGTGTGTTGGAGGCAGGGGACAGATATTAAGCCTGGAGGTAAGGGCCTGGGTAGGGAGGTG 4020

Qy 4021 TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTGGGTAGGGAGGTG 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTGGGTAGGGAGGTG 4080

Qy 4081 GTGTGAGACGCTCTCTGTCTCCTCTGCATGCCCTGAGCCCTTTGGGGAGGAGGAATGT 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4081 GTGTGAGACGCTCTCTGTCTCCTCTGCATGCCCTGAGCCCTTTGGGGAGGAGGAATGT 4140

Qy 4141 GCCCAAGGACTAAAAAAAGGCCATGGAGCCAGGGGCAACAGACCTTTCATGG 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4141 GCCCAAGGACTAAAAAAAGGCCATGGAGCCAGGGGCAACAGACCTTTCATGG 4200

Qy 4201 GCAAACCTTGGGCCCGTAGTGACCGATTGACGATGCCAATCGCATCGATACCCTTCTTC 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4201 GCAAACCTTGGGCCCGTAGTGACCGATTGACGATGCCAATCGCATCGATACCCTTCTTC 4260

Qy 4261 TTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGA 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4261 TTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGA 4320

Qy 4321 AAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCC 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4321 AAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCC 4380

Qy 4381 CCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAA 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4381 CCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAA 4440

Qy 4441 AAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAGTGAAGTCGAGTTTA 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy 2281 CTTCCTTCTTGTCATCTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281 CTTCCTTCTTGTCATCTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340

Qy 2341 GCCCCCACTACCCATAGGACAGTTGGCCTGAGTTGGCATTCGAGTTCCCTGAGACACAGAG 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341 GCCCCCACTACCCATAGGACAGTTGGCCTGAGTTGGCATTCGAGTTCCCTGAGACACAGAG 2400

Qy 2401 CCTGGCCTTTGTCTTTGGACCTGACCCAAGCTGACCCAATGTTCTCAGTACCTTTGAAT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401 CCTGGCCTTTGTCTTTGGACCTGACCCAAGCTGACCCAATGTTCTCAGTACCTTTGAAT 2460

Qy 2461 GCCCTCAAGAGCTTGAGAACCACCAGGCAGTGACATATTAGGCTAACCTGCTAACCTGGAGCT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461 GCCCTCAAGAGCTTGAGAACCACCAGGCAGTGACATATTAGGCTAACCTGCTAACCTGGAGCT 2520

Qy 2521 TGCACAGAGGCCTCAAGTGACCTCCAGGGACACAGCTGCAGAGTGGCCTTTATCC 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521 TGCACAGAGGCCTCAAGTGACCTCCAGGGACACAGCTGCAGAGTGGCCTTTATCC 2580

Qy 2581 CCAAAGAGCAACCATTTGGCATAGGTGCTGCATAGTGGCTGCAAATGGAATGCAAGTTGAATCAGTC 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581 CCAAAGAGCAACCATTTGGCATAGGTGCTGCATAGTGGCTGCAAATGGAATGCAAGTTGAATCAGTC 2640

Qy 2641 CCTTCAAGAATACTGCATGCAAGACCTAAGACTTAGAGAGAGGGGTATGCTCCTGCCC 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641 CCTTCAAGAATACTGCATGCAAGACCTAAGACTTAGAGAGAGGGGTATGCTCCTGCCC 2700

Qy 2701 CCACCCCATAAGGGGAGTGAACTATCCTAGGGGCTGGCGACCTTGGGGACGACACCAC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701 CCACCCCATAAGGGGAGTGAACTATCCTAGGGGCTGGCGACCTTGGGGACGACACCAC 2760

Qy 2761 ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCTGTCCTGCGCCCACCTCCACAC 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761 ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCTGTCCTGCGCCCACCTCCACAC 2820

Qy 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGTGGTGGACGAGGGAGGAGAGTGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGTGGTGGACGAGGGAGGAGAGTGT 2880

Qy 2881 TCCTGGGTGTGAGGGTGTAGGGGAAAGCCAGACGAGGGGAGTCTGGCTTTGTCTCCTGAA 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881 TCCTGGGTGTGAGGGTGTAGGGGAAAGCCAGACGAGGGGAGTCTGGCTTTGTCTCCTGAA 2940

Qy 2941 CACAATGTCTACTTAGTTATAACAGGCATGAGCCTCTAAAGACCCAACATCTACGACCTC 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941 CACAATGTCTACTTAGTTATAACAGGCATGAGCCTCTAAAGACCCAACATCTACGACCTC 3000

Qy 3001 TGAAAAGACAGCACCCCCACTTATACCCCTTCAAAGCAGCACTCGACACCTGCCTCTCTGAGCCTTGTGATGTG 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TGAAAAGACAGCACCCCCACTTATACCCCTTCAAAGCAGCACTCGACACCTGCCTCTCTGAGCCTTGTGATGTG 3060

Qy 3061 CCACAAAGGAGGGCATGAGTGTGAGTATAAGGCCCCAGGAGCCTTAGGAGAGGGCACTTG 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 CCACAAAGGAGGGCATGAGTGTGAGTATAAGGCCCCAGGAGCCTTAGGAGAGGGCACTTG 3120

Qy 3121 GGAAGGGGTCAGTCTGCAGAGCCCTATCCATGGAATCTGGAGCCTGGGGGTCCAACTGGTG 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 GGAAGGGGTCAGTCTGCAGAGCCCTATCCATGGAATCTGGAGCCTGGGGGTCCAACTGGTG 3180

Qy 3181 TAAATCTCTGGGCCTGCCAGGCATTCAAAGCAGCACTTATACCCTTCTCCCTGAGGATTAAACACCT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 TAAATCTCTGGGCCTGCCAGGCATTCAAAGCAGCACTTATACCCTTCTCCCTGAGGATTAAACACCT 3240

Qy 3241 GGCGGAGGGAGGCAACCCCCCACTTATACCCCTTCTCCCTCAGCCCCAGGAGGGGAGGGTAAAAA 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GGCGGAGGGAGGCAACCCCCCACTTATACCCCTTCTCCCTCAGCCCCAGGAGGGGAGGGTAAAAA 3300

Qy 3301 CTGGCCTTTCCCCCTTCCCCCTCCCATCAGGAGGAGGTTGCAGAGGGTTGCAGAGGAGGGTTGCAGAGGAA 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CTGGCCTTTCCCCCTTCCCCCTCCCATCAGGAGGAGGTTGCAGAGGGTTGCAGAGGAGGGTTGCAGAGGAA 3360

```
Db   4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAGTGAAAGTCGAGTTTA 4500

Qy   4501 CCACTCCCTATCAGTGATGATAGAGAAAGTGAAAGTGAGTTTACCACTCCCTATCAGTGAT 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 CCACTCCCTATCAGTGATGATAGAGAAAGTGAAAGTGAGTTTACCACTCCCTATCAGTGAT 4560

Qy   4561 AGAGAAAGTGAAAGTCGAGCTCGGTACCAGCAGAGACTCGGTACCAGCAGCA 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 AGAGAAAGTGAAAGTCGAGCTCGGTACCAGCAGAGACTCGGTACCAGCAGCA 4620

Qy   4621 TATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGTCAGATTTCTCCAACCA 4680
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 TATGGGATGGGATATAAAGGGGCTGGAGCACTGAGAGTCAGATTTCTCCAACCA 4680

Qy   4681 GGTAAGAGAGGAGTTCGGGTGGGGGCGTCTTCACCCACCACCAGACCTCTCCCCACCTAGAA 4740
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 GGTAAGAGAGGAGTTCGGGTGGGGGCGTCTTCACCCACCACCAGACCTCTCCCCACCTAGAA 4740

Qy   4741 GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCTGACAGGGTGGCCTT 4800
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGAGATCTGACAGGGTGGCCTT 4800

Qy   4801 CCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGTTTCCACAAGAAACACTGATGCCCCT 4860
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 CCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGTTTCCACAAGAAACACTGATGCCCCT 4860

Qy   4861 TCCCTTACGCTGTCTCTCCATCTTCCTCCTGGGATGCTCCTCCCCCGTCTTGGTTATC 4920
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 TCCCTTACGCTGTCTCTCCATCTTCCTCCTGGGATGCTCCTCCCCCGTCTTGGTTATC 4920

Qy   4921 TTGGCTCTTCGTCTTCAGCAAGATTTGCCTGTCTTCCTCTGTGTCTGTGCTCCATCCATCCATCTTCTCTACTGT 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 TTGGCTCTTCGTCTTCAGCAAGATTTGCCTGTCTTCCTCTGTGTCTGTGCTCCATCCATCCATCTTCTCTACTGT 4980

Qy   4981 CTCCGTGCTGCCTTGCCTTGCGTGTCCTTCCTTCCTTCCTTCCACCCATTTCTCACTTCACC 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 CTCCGTGCTGCCTTGCCTTGCGTGTCCTTCCTTCCTTCCTTCCACCCATTTCTCACTTCACC 5040

Qy   5041 TTTTCTCCCCTTCCTTCCTTCATTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 TTTTCTCCCCTTCTTCCTTCCTTCTCATTTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTC 5100
```

RESULT 2

US-10-024-066-5

; Sequence 5, Application US/10024066
; Publication No. US20020166134A1
; GENERAL INFORMATION:
;   APPLICANT: Field, Loren J.
;   APPLICANT: Pasumarthi, Kishore Babu S.
;   TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
;   TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
;   FILE REFERENCE: 7037-450
;   CURRENT APPLICATION NUMBER: US/10/024,066
;   CURRENT FILING DATE: 2001-12-18
;   PRIOR APPLICATION NUMBER: 60/139,942
;   PRIOR FILING DATE: 1999-06-18
;   PRIOR APPLICATION NUMBER: PCT/US00/16827
;   PRIOR FILING DATE: 2000-06-19
;   NUMBER OF SEQ ID NOS: 8
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 5
;   LENGTH: 5443
;   TYPE: DNA
;   ORGANISM: Mus musculus

US-10-024-066-5

Query Match   88.0%;   Score 5045;   DB 13;   Length 5443;
Best Local Similarity 93.9%;   Pred. No. 0;
Matches 5383;   Conservative 0;   Mismatches 60;   Indels 292;   Gaps 1;

```
Qy    1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCACCAGGTGCCCTAGTCTCAATTTCAGT 60
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCACCAGGTGCCCTAGTCTCAATTTCAGT 60

Qy   61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCGGCCTCGGCCTGGCTAATTTGGACTTTGTTTTTAT 120
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCGGCCTCGGCCTGGCTAATTTGGACTTTGTTTTTAT 120

Qy  121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTCACCTAGGGTGCCCGTGA 180
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTCACCTAGGGTGCCCGTGA 180

Qy  181 AGCCCTCAGACCTGGAGCCCTTGCAACAGCCCTTAGTGGAAGCAGAATAAAGCAATTT 240
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 AGCCCTCAGACCTGGAGCCCTTGCAACAGCCCTTAGTGGAAGCAGAATAAAGCAATTT 240

Qy  241 TCCTTAAAGCCAAAATCCTGCCTCAGCTGTGGTCCAC 300
        |||||||||||||||||||||||||||||||||||||
Db  241 TCCTTAAAGCCAAAATCCTGCCTCAGCTGTGGTCCAC 300

Qy  301 AGGGTGGGGAGGTGGGGCTTGAGAAGGTGGGAGTGGCAAAAGCCGATCCCTAG 360
        |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  301 AGGGTGGGGAGGTGGGGCTTGAGAAGGTGGGAGTGGCAAAAGCCGATCCCTAG 360

Qy  361 GGCCCGTGTGAAGTCGGGAGCCTTCGGAGCTGCTCATAGATCCTCCAGCC 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  361 GGCCCGTGTGAAGTCGGGAGCCTTCGGAGCTGCTCATAGATCCTCCAGCC 420

Qy  421 AAACATAGCAAGAGAGTGATACCTCCTTTGTGACTTCCCCAGGCCGTACCTGTCAGGTT 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  421 AAACATAGCAAGAGAGTGATACCTCCTTTGTGACTTCCCCAGGCCGTACCTGTCAGGTT 480
```

```
Qy   5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCC 5160
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCC 5160

Qy   5161 CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTTATGTAAACAATCTTCCA 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTTATGTAAACAATCTTCCA 5220

Qy   5221 GTGAGCCACAGCTTCAGTGGGGTGCTGCTGGGTGCTCCTCCTTACCTTCTCCACCCCTGGCTGTC 5280
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 GTGAGCCACAGCTTCAGTGGGGTGCTGCTGGGTGCTCCTCCTTACCTTCTCACCCCCTGGCTGTC 5280

Qy   5281 CTGTTCCATCCTGTCAGGATCTTCTAGATTGGTCTCCCAGCCTTCCAGCCTCCTCTCTCCT 5340
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5281 CTGTTCCATCCTGTCAGGATCTTCTAGATTGGTCTCCCAGCCTTCCAGCCTCCTCTCTCCT 5340

Qy   5341 GCCTGTTCCTCTCTGTCCAGCTGCGCCACTTGGGTGCCTGGTGCCTCGTTCCAGCTGGTGGTCCAC 5400
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5341 GCCTGTTCCTCTCTGTCCAGCTGCGCCACTTGGGTGCCTGGTGCCTCGTTCCAGCTGGTGGTCCAC 5400

Qy   5401 ATTCTTCAGGATTCTCTGAAAAGTTAACCAGGTGAGAATGTTCCCCGTGCCCTGTAGACAGCAGA 5460
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5401 ATTCTTCAGGATTCTCTGAAAAGTTAACCAGGTGAGAATGTTCCCCGTGCCCTGTAGACAGCAGA 5460

Qy   5461 TCACGATTCTCCCGGAAGTCAGGCTTCAGCCTTCCCCGGCCTGCCCAGCTGCTCCTCCAGCC 5520
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5461 TCACGATTCTCCCGGAAGTCAGGCTTCAGCCTTCCCCGGCCTGCCCAGCTGCTCCTCCAGCC 5520

Qy   5521 CTCTTAGCAAACCTTCAGGCACCCTTACCCCACACATAGACCTCTGACAGAGAAGCAGGCACT 5580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5521 CTCTTAGCAAACCTTCAGGCACCCTTACCCCACACATAGACCTCTTTGTGACTTACCTGTCAGGTT 5580
```

```
Qy  1561 AGCCAAGAGACACCCGTAGGGAGGAGGATCACACTGGATGAAGAGGAGATGTGTGGAGAAGTCCAGG 1620
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561 AGCCAAGAGACACCCGTAGGGAGGAGGATCACACTGGATGAAGAGGAGATGTGTGGAGAAGTCCAGG 1620

Qy  1621 GCAACCTAGGAGCCAGAGCCTAAAAGAGCAAGAGAGATAAAGGTGCTTCAAAGGTGGCCAGG 1680
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621 GCAACCTAGGAGCCAGAGCCTAAAAGAGCAAGAGAGATAAAGGTGCTTCAAAGGTGGCCAGG 1680

Qy  1681 CTGTGCACACAGAGGGTCGAGGACTGGTCGTGTAGAGCCTCAAGATAAGGATGATGCTCAGA 1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681 CTGTGCACACAGAGGGTCGAGGACTGGTCGTGTAGAGCCTCAAGATAAGGATGATGCTCAGA 1740

Qy  1741 ATGGCGGGGGGGGGGATTCTGGGGGGGGGAGAGAGGAGCCTGAGAACAG 1800
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741 ATGGCGGGGGGGGGGATTCTGGGGGGGGGAGAGAGGAGCCTGAGAACAG 1800

Qy  1801 AGAATCTGAAGCGCTGGAAACGATACCATAAAGGAACAACCAGGCTACCTTTAGATG 1860
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801 AGAATCTGAAGCGCTGGAAACGATACCATAAAGGAACAACCAGGCTACCTTTAGATG 1860

Qy  1861 TAAATCATGAAAGACAGGGAAGCTGAGAGCTGAGAGAGCTAGAGGACCCCGGGGCAAGA 1920
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861 TAAATCATGAAAGACAGGGAAGCTGAGAGCTGAGAGAGCTAGAGGACCCCGGGGCAAGA 1920

Qy  1921 CATGGAAGCAAGAACAAGCCAGGTTGAGCGTCCGTGAAATCAGCCTGCTGAAGGCAGAG 1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921 CATGGAAGCAAGAACAAGCCAGGTTGAGCGTCCGTGAAATCAGCCTGCTGAAGGCAGAG 1980

Qy  1981 CCCTGGTATGAGCACCAGAACAGCAGAGGTAGGGTTAATGTCGAGACAGGGAACAGAAG 2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981 CCCTGGTATGAGCACCAGAACAGCAGAGGTAGGGTTAATGTCGAGACAGGGAACAGAAG 2040

Qy  2041 GTAGACAGAGGAACAGACAGAGACCGGGGAGCCAGTGAACAAAAGGAATGGTCCTTCTCAC 2100
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041 GTAGACAGAGGAACAGACAGAGACCGGGGAGCCAGTGAACAAAAGGAATGGTCCTTCTCAC 2100

Qy  2101 CTGTTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGACTGTTCATCAGACTGCAGGGC 2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101 CTGTTGGCCAGAGCGTCCATCTGTGTCCACATACTCTAGACTGTTCATCAGACTGCAGGGC 2160

Qy  2161 TGGCTTGGAGGGCAGCTGGAAAGTATGTGAGAGCAAGGGGGAGACAAGGGGCCTAGGA 2220
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161 TGGCTTGGAGGGCAGCTGGAAAGTATGTGAGAGCAAGGGGGAGACAAGGGGCCTAGGA 2220

Qy  2221 AAGGAAGAGAGGGCAAACCAGGCCACAAGAGAGGGCAGAGCCCAGAACTGAGTTAACTC 2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221 AAGGAAGAGAGGGCAAACCAGGCCACAAGAGAGGGCAGAGCCCAGAACTGAGTTAACTC 2280

Qy  2281 CTTCCTTGTCATCTTCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281 CTTCCTTGTCATCTTCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA 2340

Qy  2341 GCCCCACTACCAAGTTTGGCTGAGTGCATTCTAGGTTCCCTGAGGACAGAG 2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341 GCCCCACTACCAAGTTTGGCTGAGTGCATTCTAGGTTCCCTGAGGACAGAG 2400

Qy  2401 CCTGCCCTTTGTCTCTTGGACCTGACCTGGGAACTCTGTGACCACCATCCCCCATGGCTAACCCTGGAGCT 2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401 CCTGCCCTTTGTCTCTTGGACCTGACCTGGGAACTCTGTGACCACCATCCCCCATGGCTAACCCTGGAGCT 2460

Qy  2461 GCCCTCAAGAGCTTGAGACACCAGGCAGTGACCATATTAGGCTGACATATTAGGCTAACCCTGGAGCT 2520
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461 GCCCTCAAGAGCTTGAGACACCAGGCAGTGACCATATTAGGCTGACATATTAGGCTAACCCTGGAGCT 2520

Qy  2521 TGCACACAGGAGCTCAAGTGACCTCCAGGGACACAGCTGCAGCAGTGGCCTTTATCC 2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521 TGCACACAGGAGCTCAAGTGACCTCCAGGGACACAGCTGCAGCAGTGGCCTTTATCC 2580

Qy  2581 CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGAATGGGATTGAATCAGGTC 2640
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581 CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGAATGGGATTGAATCAGGTC 2640

Qy  2641 CCTTCAAGAATACTGCATGCAAGCACCCTAGACCCCTGAGAGAGGGGTATGTCCTGCCC 2700
```

```
Qy   481 GAAACAGATTTAGAGAGCCTCACCTGAACTCTGAAGCTCATCCACCAAGCA 540
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   481 GAAACAGATTTAGAGAGCCTCACCTGAACTCTGAAGCTCATCCACCAAGCA 540

Qy   541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCGTGATAATATGCGAGAGTGGGCCAC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCGTGATAATATGCGAGAGTGGGCCAC 600

Qy   601 AGAAGTCCTGGGGTGTAGGACTGACCAGTGACTTTTCAGTCGGCAAAGTATGACCCCC 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   601 AGAAGTCCTGGGGTGTAGGACTGACCAGTGACTTTTCAGTCGGCAAAGTATGACCCCC 660

Qy   661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCCATCCAGGCAGTCTCTAAGAGGACATG 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCCATCCAGGCAGTCTCTAAGAGGACATG 720

Qy   721 GGATGAGAGATGTAGTCATGTGCATTCCAAACACAGCTATCCACAGTGTCCCCTTGCCCC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   721 GGATGAGAGATGTAGTCATGTGCATTCCAAACACAGCTATCCACAGTGTCCCCTTGCCCC 780

Qy   781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTCTTCCTCCCCATCCTCCCAG 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTCTTCCTCCCCATCCTCCCAG 840

Qy   841 GACACACCCCTGGTCTTGCAGTATTCATTTCTTCCTTCACGTCCCCTGTGACTTCCAT 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   841 GACACACCCCTGGTCTTGCAGTATTCATTTCTTCCTTCACGTCCCCTGTGACTTCCAT 900

Qy   901 TTTGCAAGGCTTTTGACCTTGCTCGAGCTGGCTGGCCCTAGTGTGTGCAA 960
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   901 TTTGCAAGGCTTTTGACCTTGCTCGAGCTGGCTGGCCCTAGTGTGTGCAA 960

Qy   961 GCCATCTCAAGAGAGAAAGCAGACAACACAGGGGGACCCAGATTTTGGAAGGATCAGGAACTAAA 1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   961 GCCATCTCAAGAGAGAAAGCAGACAACACAGGGGGACCCAGATTTTGGAAGGATCAGGAACTAAA 1020

Qy  1021 TCACTGCGGGGGCCTGGGGTAGAAAAAGAGTGAGTGACCCTCAAGCTAAGCCAAGC 1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1021 TCACTGCGGGGGCCTGGGGTAGAAAAAGAGTGAGTGACCCTCAAGCTAAGCCAAGC 1080

Qy  1081 TAGTCCCCGAGATACTCTGCCCACAGCTCGGGGGTAGCTTTAGGAATGTGGGTC 1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1081 TAGTCCCCGAGATACTCTGCCCACAGCTCGGGGGTAGCTTTAGGAATGTGGGTC 1140

Qy  1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACTACAGAC 1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1141 TGAAAGACAATGGGATTGGAAGACATCTCTTTGAGTCTCCCCTCAACCCCACTACAGAC 1200

Qy  1201 ACACTCGTGTGTGCCAGACTCCTGTTCAACAGCCCTGTGTTGTCTGACCACTGAGCTAG 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201 ACACTCGTGTGTGCCAGACTCCTGTTCAACAGCCCTGTGTTGTCTGACCACTGAGCTAG 1260

Qy  1261 GCAACCAGGAGCATGGGCCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAACAG 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261 GCAACCAGGAGCATGGGCCCCTGTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAACAG 1320

Qy  1321 CAGGGGAGGGAGAACAGAGAACAGAGAATAAGGAAGAAGAGGCCAGTCAATCAGA 1380
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321 CAGGGGAGGGAGAACAGAGAACAGAGAATAAGGAAGAAGAGGCCAGTCAATCAGA 1380

Qy  1381 TGCAGTCAGAAGAGCATGGGCCCCTGTGCTGAGCAGCAGAGGAAACAGAAAAGGGAG 1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381 TGCAGTCAGAAGAGCATGGGCCCCTGTGCTGAGCAGCAGAGGAAACAGAAAAGGGAG 1440

Qy  1441 AGATTCTGGCATAAGGAGGCCACAGAAGAGAAGAGCCCCAGTCTCCTTT 1500
         |||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441 AGATTCTGGCATAAGGAGGCCACAGAAGAGAAGAGCCCCAGTCTCCTTT 1500

Qy  1501 ATACCCTCATCCCGTCTCCCCAATTAAGCCACTCTTCCTTCCTCCCACTCTTCCTTCTTCCTAGATCAGACCTGAGCTGC 1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501 ATACCCTCATCCCGTCTCCCCAATTAAGCCACTCTTCCTTCCTCCCACTCTTCCTTCTTCCTAGATCAGACCTGAGCTGC 1560
```

```
Db  2641  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  2700
          CCTTCAAGAAATACTGCATGCAAGACCTTAAGACCCCTGAGAGACGGGGTATGCTCCTGCCC
Qy  3721  TACAGACAGTATCTGTACAGAGTAAGGAGGAGAGAGAGGGGTGTAGAATTCTCTTACTA  3780

Db  2701  CCACCCACCATAAGGGAGTGAACTATCCTAGGGGGCTGGCGACCTGGGAGACACCAC  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3781  TCAAAGGGAAACTGAGTCGTGCACCTGGACCTGCACCTGCAGCACATCATGACTT  3840

Db  2761  ATTACTGAGAGTGCTGAGCCCAGAAAAACTGACCGCCCTGTCCTGCCCACCTCCACAC  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3841  TGTCTCTGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGAGTAGGAGGCTCCCCTCA  3900

Db  2821  TCCTAGAGCTATATTGAGAGTGCAGTAGATAGGGTGGAAGCTGGCAGGGAGAGTGT  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3901  GCCTGAAGCTATGCAGATAGCCAGGGTTGAAAGGGGAAGGGGAGAGCCTGGGATGGGAGC  3960

Db  2881  TCCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCCAGGGAGTCTGGCTTTGTCTCCTGAA  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  3961  TTGTGTGTTGGAGGCAGGGGACAGATATTAAGGCTGAAGAGAAGGTGACCCTTACCCAG  4020

Db  2941  CACAATGTCTACTTAGTTATAACAGGCATGACCTGCTAAAGACCCAACATCTACGACCTC  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4021  TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTTGGGGAGGAGGAATGT  4080

Db  3001  TGAAAAGACAGCAGCCCTGGAGGACACAGGGGTTGTCTCTGAGCCTTGGGTGCTTGATGGTG  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4081  GTGTGAGACGCTCCGTCTCTCCTCTGCGTCCCATGCCCTTTGGGGAGGCGGAGGGAATGT  4140

Db  3061  CCACAAAGGAGGGCATGAGTGTGAGTATAAGGCCACTTG  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4141  GCCCAAGGACTAAAAAAAAGGCCATGGACCCAGGGGCCAGGAGGGCAACAGACCTTTCATGG  4200

Db  3121  GGAAGGGGTCAGTCTGCGAGAGCCCTATCCATGGAATCTGGAGCCTGGGGCCAACTGGTG  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4201  GCAAACCTTGGGCCGTAGTGATCGATTGACAACGAACTCGCCAATCGATACCCTTCTC  4260

Db  3181  TAAAATCTTCTCGGGCCTGCCAGGCATTCAAAAGCACACCACCTCTGCGCCCCAGGATTAAACACTT  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4261  TTCTAACGGACGAGGAGGAACTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAGTGA  4320

Db  3241  GGCGGAAGGGAGCAACCCCCCCCCCCACTTATACCCTTTCTCCCCTCAGCGCCCCAGGATTAAACACCT  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4321  GGAGCCAGGACAGGAGGGAAGTGGAGGGAG  4291

Db  3301  CTGGCCTTCCCCCCTTCCCCACCTCCCATCCAGGAGTGCAGAGGGTTGCAGAGGGGAGGGTAAAAA  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4292  AAGTCGAGTTTACCACTCCCTATCAGTGATGAAAAGTGAAAGTGAAAAGTCGAGTTTACCACTC  4380

Db  3361  CCTACACATGTCAAAACATCATGTGCACGAATCATGTGTGTAAGTCTGTGTGCATGTGTAAAGTCTGTG  3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4381  CCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAA  4440

Db  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTAAAGTCTGTGTGCATGTGTGTCTG  3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4441  AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA  4500

Db  3481  ACTGAAAACGGGCATGCGTGTCCAGCTGTTCAGTTCTGTGCGTGAGCTGTTACCAGACTGCA  3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4501  CCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGAT  4560

Db  3541  GGTTTGTGTGTAAATTGCCCAGCTGTAAATGGGTGAATCCCATGGTCTCTATATGTTGTCT  3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4561  AGAGAAAGTGAAAGTCGAGTTACCAGCAGGACTCCAAATTTAGGCAGCAGGCA  4620

Db  3601  TGGATGATGGCCTGCATCTCAAGGACCATGGGTGAATCCCATGGTCTCTATATGTGTCT  3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4292  GGTCCCAGCAGGACTCCAAATTTAGGCAGCAGGCA  4328

Db  3661  CTAAGCTAAGGTAGCAGGTCTTTGGGACGACACCTGTCTAGAGATGTGGGCAACAGAGAC  3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4621  TATGGGATGGATATAAAGGGGCTGGAGCACTGAGAGATTTCTCCAACCCA  4680

Db  3721  TACAGACAGTATCTGTACAGAGTAAGGAGGAGAGAGAGGGGGTGTAGAATTCTCTTACTA  3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4329  TATGGGATGGATATAAAGGGGCTGGAGCACTGAGAGATTTCTCCAACCCA  4388

Db        GGTAAGAGGGAGTTTCGGGTGGGGTGGGGTGGGGTGGGGGTCGGGATCTGAAGTCTGACAGAGAC  4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4681  GGTAAGAGGGAGTTTCGGGTGGGGTGGGGTGGGGGTCGGGATCTGGAAGTCTGACAGGTGGCTT  4800

Db  4441  GGAAACTGCCTTTCCTGGAAGTGGGTTCAGGCGGGTTCAGGCGGGTTCAGGTGGCCTT  4508
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy  4801  CCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAACACTGGATGCCCCT  4860

Db  4509  CCACCAGCCTGGGAAGTTCTCAGTGGCAGGAGGTTTCCACAAGAACACTGGATGCCCCT  4568
```

; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/JP01/06086
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: JP2000-215143
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
; LENGTH: 5443
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(5443)
US-10-332-966-2

Query Match   88.0%; Score 5045; DB 16; Length 5443;
Best Local Similarity   93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 0; Indels 60; Gaps 1;

```
Qy      1 GGATCCTGCAAGGTCACACAAGGGTCTCACCCACCAGGTGCCCTAGTCTCCAATTTCAGT   60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCCACCAGGTGCCCTAGTCTCCAATTTCAGT   60

Qy     61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCAGAGCTAATTTGGACTTTGTTTTTAT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCAGAGCTAATTTGGACTTTGTTTTTAT  120

Qy    121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTCTACCCAGCTCTAAGGGTGCCCGTGA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TTCAAAAGGGCCTGAATGAGGAGTAGATCTTGTGTCTACCCAGCTCTAAGGGTGCCCGTGA  180

Qy    181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTAGGTGGAAGCAGAATAAAGCAATTT  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 AGCCCTCAGACCTGGAGCCTTTGCAACAGCCCTTAGGTGGAAGCAGAATAAAGCAATTT  240

Qy    241 TCCTTAAAGCCAAAATCCTGCCCTCCTTAGACTCTTCTCTCTGACCTCGGTCCTGGGCTCT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TCCTTAAAGCCAAAATCCTGCCCTCCTTAGACTCTTCTCTCTGACCTCGGTCCTGGGCTCT  300

Qy    301 AGGGTGGGAGGTGGGGCTTGGAAGAAGAGAGGTGGGGAAGGTGGCAAAAGCCGATCCCTAG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGGGTGGGAGGTGGGGCTTGGAAGAAGAGAGGTGGGGAAGGTGGCAAAAGCCGATCCCTAG  360

Qy    361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCTCCAGCC  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCTCCAGCC  420

Qy    421 AAACATAGCAAGAAGTGATACCTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AAACATAGCAAGAAGTGATACCTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT  480

Qy    481 GAAACAGGATTTAGAGAAGCCTCTGAACTCACCTGAACTCTGAAGCTCATCCACCAAGCA  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GAAACAGGATTTAGAGAAGCCTCTGAACTCACCTGAACTCTGAAGCTCATCCACCAAGCA  540

Qy    541 AGCACCTAGGTGCCACTGTCTAGTAGTATCCTAGCGCTGATAATATGCAGAATGCGGGCCAC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGCACCTAGGTGCCACTGTCTAGTAGTATCCTAGCGCTGATAATATGCAGAATGCGGGCCAC  600

Qy    601 AGAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCGGCAAAGGTATGACCCCC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 AGAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCGGCAAAGGTATGACCCCC  660

Qy    661 TCAGCAGATGAGTAGTAATGTCCCCTTAGATCCCATCCCAGGCAGGTCTCTAAGAGGACATG  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCAGCAGATGAGTAGTAATGTCCCCTTAGATCCCATCCCAGGCAGGTCTCTAAGAGGACATG  720

Qy    721 GGATGAGAGATGTAGTCATGTGGCATTCCAAACAGCAGTCCTTGCCCC  780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GGATGAGAGATGTAGTCATGTGGCATTCCAAACAGCAGTCCTTGCCCC  780

Qy    781 TTCCACTTAGCCAGGAGGACAGTAACCTTAGCCTATCTTTCTTCCCCATCCTCCCAG  840
```

```
Qy   4861 TCCCTTACGCTGCTTCCATCTTCCTCCTGGGGATGCTCCTCCCCCGTCTTCGTTTATC  4920
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4569 TCCCTTACGCTGCTTCCATCTTCCTCCTGGGGATGCTCCTCCCCCGTCTTCGTTTATC  4628

Qy   4921 TTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGCCTGCCACTCCACTCCATCTTTCTACTGT  4980
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4629 TTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTGCCTGCCACTCCACTCCATCTTTCTACTGT  4688

Qy   4981 CTCCGTGCCTTGCCTTCTTGGCTGTCCTTCCACCCATTTCTCACTCACC  5040
            |||||||||||||||||||||||||||||||||||||||||||||||||
Db   4689 CTCCGTGCCTTGCCTTCTTGCGTGTCCTTCCCACCCATTTCTCACTCACC  4748

Qy   5041 TTTTCTCCCCTTCTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTC  5100
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4749 TTTTCTCCCCTTCTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTC  4808

Qy   5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  5160
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4809 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  4868

Qy   5161 CTGTGTCGAGAGTGCTGAGAATCACACACCTGGGGTTCCACCCCTTATGTAAACAATCTTCCA  5220
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4869 CTGTGTCGAGAGTGCTGAGAATCACACACCTGGGGTTCCACCCCTTATGTAAACAATCTTCCA  4928

Qy   5221 GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTCTTACCTTCTCCACCCCTGCCTTGTC  5280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4929 GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTCTTACCTTCCTCACCCCTGCCTTGTC  4988

Qy   5281 CTGTTCCATCCTGGTCAGGATCCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCTCCT  5340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4989 CTGTTCCATCCTGGTCAGGATCCTAGATTGGTCTCCCAGCCTCTGCTACTCCTCTCCT  5048

Qy   5341 GCCTGTTCCTCTCTCGTCCAGCTGCGCCACTGGTGCCTGCCACTGTGGTCCAC  5400
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5049 GCCTGTTCCTCTCTCGTCCAGCTGCGCCACTGGTGCCTGCCACTGTGGTCCAC  5108

Qy   5401 ATTCTTCAGGATTCTGAAAAGTAACCAGGTGAGAGATGTTCCCCGTGTAGACAGAGA  5460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5109 ATTCTTCAGGATTCTGAAAAGTAACCAGGTGAGAGATGTTCCCCGTGTAGACAGAGA  5168

Qy   5461 TCACGATTCTCCCGGAAGCTCAGGCTTCCAGCCCTCTCTTCTCTGCGCCAGCTGCCCGGCA  5520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5169 TCACGATTCTCCCGGAAGCTCAGGCTCAGGCTTCCAGCCCTCTCTTCTCTGCGCCAGCTGCCCGGCA  5228

Qy   5521 CTCTTAGCAAACCTCAGGCACCCTTACCCCACAATAGACCTCTGACAGAGAAGCAGGCACT  5580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5229 CTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTCTGACAGAGAAGCAGGCACT  5288

Qy   5581 TTACATGGAGTCTGGTGGGAGCCCATAGGCTACGGTGTAAAAGAGGAGGAAGTGGT  5640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5289 TTACATGGAGTCTGGTGGGAGCCCATAGGCTACGGTGTAAAAGAGGAGGAAGTGGT  5348

Qy   5641 GGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACACCAGAAATGACAGAGACA  5700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5349 GGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACACCAGAAATGACAGAGACA  5408

Qy   5701 TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC  5735
            ||||||||||||||||||||||||||||||||||
Db   5409 TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC  5443
```

RESULT 3
US-10-332-966-2
; Sequence 2, Application US/10332966
; Publication No. US20030188324A1
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, Koji
; APPLICANT: KAWASE, Yosuke
; APPLICANT: SUZUKI, Hiroshi
; TITLE OF INVENTION: p300 TRANSGENIC ANIMAL
; FILE REFERENCE: 382.1040
; CURRENT APPLICATION NUMBER: US/10/332,966

```
Db  781  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         TTCCACTTAGCCAGGAGACAGTAACCTTAGCCTATCTTTCTTCCTCCCCATCCTCCCAG   840

Qy  841  GACACACCCCTGGTCTGCAGTATTCATTTCTTCACGTCCCCTCGTTGACTTCCAT        900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  841  GACACACCCCTGGTCTGCAGTATTCATTTCTTCACGTCCCTCGTTGACTTCCAT         900

Qy  901  TTGCAAGGCTTTTGACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA   960
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  901  TTGCAAGGCTTTTGACCTCTGCAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA   960

Qy  961  GCCATCTCAAGAGAAAAGCAGACAGAACAGGGGACCAGAGATTTTGAAGGATCAGGAACTAAA   1020
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  961  GCCATCTCAAGAGAAAAGCAGACAGAACAGGGGACCAGAGATTTTGAAGGATCAGGAACTAAA   1020

Qy 1021  TCACTGGCGGGGCCTGGGGTAGAAAAGAGTGAGTGCTCCGCTCAAGCTAAGCCAAGC       1080
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1021  TCACTGGCGGGGCCTGGGGTAGAAAAGAGTGAGTGCTCCGCTCAAGCTAAGCCAAGC       1080

Qy 1081  TAGTCCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC   1140
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1081  TAGTCCCCGAGATACTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGGTC   1140

Qy 1141  TGAAAAGACAAATGGAATGGGAAGACATCTCTTTGAGTCTCCCCCTCAACCCCACTACGAGAC   1200
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1141  TGAAAAGACAAATGGAATGGGAAGACATCTCTTTGAGTCTCCCCCTCAACCCCACTACGAGAC   1200

Qy 1201  ACACTCGTGTGTGGCCAGACTCCTCTGTGTTCTGACCACTGAGCTAG                 1260
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db 1201  ACACTCGTGTGTGGCCAGACTCCTCTGTGTTCTGACCACTGAGCTAG                 1260

Qy 1261  GCAACCAGAGCATGGGCCCTGTGTCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG   1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1261  GCAACCAGAGCATGGGCCCTGTGTCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG   1320

Qy 1321  CAGGGGGAGGGAGAACAGAGAACGAAATAAGGAAGAAGGAAGAAAAGGCCAGTCAATCAGA   1380
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1321  CAGGGGGAGGGAGAACAGAGAACGAAATAAGGAAGAAGGAAGAAAAGGCCAGTCAATCAGA   1380

Qy 1381  TGCAGTCAGAGAGACATGGGAAGCCAACCAGAGAAACAGAAAACAGAAAACAGAAAAGGGAG   1440
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1381  TGCAGTCAGAGAGACATGGGAAGCCAACCAGAGAAACAGAAAACAGAAAACAGAAAAGGGAG   1440

Qy 1441  AGATTCTGGGAAGACATACCATAAAGAGCAAGATAAAGGTGCTTCAAAGGTGGCCAGG     1500
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1441  AGATTCTGGGAAGACATACCATAAAGAGCAAGATAAAGGTGCTTCAAAGGTGGCCAGG     1500

Qy 1501  ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTCTAGATCAGACCTGAGCTGC   1560
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1501  ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTCTAGATCAGACCTGAGCTGC   1560

Qy 1561  AGCGAAGAGAAGGGCGGGGGGGGGATTCTCGGGGGAGAGAGATCGACTGATGAAGGAAGTGTGGAACAG   1620
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1561  AGCGAAGAGAAGGGCGGGGGGGGGATTCTCGGGGGAGAGAGATCGACTGATGAAGGAAGTGTGGAACAG   1620

Qy 1621  GCAACCTTAAGAGCGCTGGGGGGAGGCCTAAAGCAAGATAAAGGTGCTTCAAAGGTGGCCAGG   1680
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1621  GCAACCTTAAGAGCGCTGGGGGGAGGCCTAAAGCAAGATAAAGGTGCTTCAAAGGTGGCCAGG   1680

Qy 1681  CTGTGCACACAGAGGGTCTGAGCCCGTGGTGGTAGAGACTGCTGGTGGTGGTAGGGGGAGAGAACCTGGAACACG   1740
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1681  CTGTGCACACAGAGGGTCTGAGCCCGTGGTGGTAGAGACTGCTGGTGGTGGTAGGGGGAGAGAACCTGGAACACG   1740

Qy 1741  ATGGGCGGGGGATTCTCGGGGGAGAGAAGAGTGACTGATGAAGGAAGTGAGAACCCAGCCTACCTTTAGATG   1800
         |||||||||||||||||||||||||||||||||||||||||||||||
Db 1741  ATGGGCGGGGGGATTCTCGGGGGAGAGAAGAGTGACTGATGAAGGAAGTGAGAACCCAGCCTACCTTTAGATG   1800

Qy 1801  AGAATCTGGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTGGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTGAAACGCGTCGAAACTCTG   1860

Db 1861  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         TAAATCATGAAAGACAGCAGGGAGAAGGGAGAAGAAGCTGGAGAGAGTAGAAGGACCCGGGGCAAGA   1920

Qy 1921  CATGGAAGCAAGGACAAGGACAAGCCAGGTTGAGCGGCTCCGTGAAATCAGCCTGCTGAAAGGCAGAG   1980
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1921  CATGGAAGCAAGGACAAGGACAAGCCAGGTTGAGCGGCTCCGTGAAATCAGCCTGCTGAAAGGCAGAG   1980

Qy 1981  CCCTGGTATGAGACAGCAGACAGCAGAGGCTAGGTTAATGTCGAGACAGGAACAGAGAAG   2040
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1981  CCCTGGTATGAGACAGCAGACAGCAGAGGCTAGGTTAATGTCGAGACAGGAACAGAGAAG   2040

Qy 2041  GTAGACACAGGAACAGACAGAGAGACGGGGAGCCAGGTAACAGTAACAAAGAAATGGTCCTTCTCAC   2100
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2041  GTAGACACAGGAACAGACAGAGAGACGGGGAGCCAGGTAACAGTAACAAAGAAATGGTCCTTCTCAC   2100

Qy 2101  CTGTGGCCAGAGCGTCCATCTGTCTCCACATACTCTAGAAATGTTCATCAGACTGCAGGGC   2160
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2101  CTGTGGCCAGAGCGTCCATCTGTCTCCACATACTCTAGAAATGTTCATCAGACTGCAGGGC   2160

Qy 2161  TGGCTTGGGAGGCAGCTGGAGAAGTATGTGAGAGCCAGGGGAGACAAGGGGCCTAGGA     2220
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2161  TGGCTTGGGAGGCAGCTGGAGAAGTATGTGAGAGCCAGGGGAGACAAGGGGCCTAGGA     2220

Qy 2221  AAGGAAGAGAGGGCAAACCAGGACAACCAGGCCACACAAGAGGGCAGAGCCCAGAACTGAGTTAACTC   2280
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2221  AAGGAAGAGAGGGCAAACCAGGACAACCAGGCCACACAAGAGGGCAGAGCCCAGAACTGAGTTAACTC   2280

Qy 2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA   2340
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGACCACCATCCCCCATGA   2340

Qy 2341  GCCCCCACTACCCATACCACCATACCACCTTCTAGGTTCCCTGAGGACAGAG           2400
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2341  GCCCCCACTACCCATACCACCATACCACCTTCTAGGTTCCCTGAGGACAGAG           2400

Qy 2401  CCTGGCCTTGTCTTCTTGACCTGACCCAAGCTGACCCAATGTTCTCAGTACCTTTGAAT   2460
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2401  CCTGGCCTTGTCTTCTTGACCTGACCCAAGCTGACCCAATGTTCTCAGTACCTTTATCAT   2460

Qy 2461  GCCCTCAAGAGCTGACTTGAGAACCAGCAGTGACATATTAGGCTACATTAGGCTACATTAGGCGTCTAACCCTGGAGCT   2520
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2461  GCCCTCAAGAGCTGACTTGAGAACCAGCAGTGACATATTAGGCTACATTAGGCTAACCCTGGAGCT   2520

Qy 2521  TGCACACAGGAGCCTCAAGTGACCTCAGGGACACAGTCTGCAGACAGAGTTGCCCTTTATCC   2580
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2521  TGCACACAGGAGCCTCAAGTGACCTCAGGGACACAGTCTGCAGACAGAGTTGCCCTTTATCC   2580

Qy 2581  CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGAATGGGAATGGGAATGGGAATGTTGAATCAGGTC   2640
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2581  CCAAAGAGCAACCATTTGGCATAGGTGGCTGCAAATGGGAATGGGAATGTTGAATCAGGTC   2640

Qy 2641  CCTTCAAGAATACTGCATGCAAGACCCCTGCCCTGGGGGGTATGCTGCCC           2700
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2641  CCTTCAAGAATACTGCATGCAAGACCCCTGCCCTGGGGGGTATGCTGCCC           2700

Qy 2701  CCACCCACCATAAGGGAGTGAACTATCCTCAGGGGGGTCGCGACCTTGGGGAGACACCAC   2760
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2701  CCACCCACCATAAGGGAGTGAACTATCCTCAGGGGGGTCGCGACCTTGGGGAGACACCAC   2760

Qy 2761  ATTACTGAGAGTGCTGAGCCCGAGCCCTCAGATAAGGAGATGATGCTTCAGA             2820
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2761  ATTACTGAGAGTGCTGAGCCCGAGCCCTGCTGAGCCCGAGCCCTGCTGACCCCACAC       2820

Qy 2821  TCTAGAGCTATATTGAGAGGTGACAGTGGGGGAGACCTGGAACAGTGCAGGGGAGAGTGT   2880
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2821  TCTAGAGCTATATTGAGAGGTGACAGTGGGGGAGATAGGGTGGCAGGGGAGAGTGT       2880

Qy 2881  TCCTGGGTGTGAGGGTGTAGGGGTAGGGAAAGCCAGCAGGGGAGTCTGGCTTTGTCTTCCTGAA   2940
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2881  TCCTGGGTGTGAGGGTGTAGGGGTAGGGAAAGCCAGCAGGGGAGTCTGGCTTTGTCTTCCTGAA   2940

Qy 2941  CACAATGTCTACTTAGTTTATAACAGGCATGAACCCAACATCTACGACCTC            3000
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2941  CACAATGTCTACTTAGTTTATAACAGGCATGAACCCAACATCTACGACCTC            3000
```

Qy 4081 GTGTGAGACGCTCCTGTCTCCTCCTGCATGCCCTTGGGGAGGAGGAATGT 4140
          |||||||||||||||||||||||||||||||| ||||||||||||||||||
Db 4081 GTGTGAGACGCTCCTGTCTCCTCCTCCTATCTGCCCATCGGCCCCTTTGGGGAGGAGGAATGT 4140

Qy 4141 GCCCAAGGACTAAAAAAGGCCATGAGCCAGGAGGGCAACAGACCTTTCATGG 4200
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4141 GCCCAAGGACTAAAAAAGGCCATGAGCCAGGAGGGCAACAGACCTTTCATGG 4200

Qy 4201 GCAAACCTTGGGCCCGTAGTGATCGATTGACAAGAACTCGCCAATCGATACCCTTCTTC 4260
          ||||||||||||| ||||||
Db 4201 GCAAACCTTGGGCCCGTGTCCTGCGTGTCGCAAGGGATCAAAGGAGGA 4260

Qy 4261 TTCTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAAGTGA 4320
Db 4261 GGAGCCAGGACAGGAGGTGGGAGGGAG-------------------- 4291

Qy 4321 AAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTC 4380
Db 4292 -------------------------------------- 4291

Qy 4381 CCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAA 4440
Db 4292 -------------------------------------- 4291

Qy 4441 AAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAGAAAAAGTGAAAGTCGAGTTTA 4500
Db 4292 -------------------------------------- 4291

Qy 4501 CCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGAT 4560
Db 4292 -------------------------------------- 4291

Qy 4561 AGAGAAAGTGAAAGTCGAGCTACCAGCAGGACTCGAGACTCGTCAGAGATTCTCCAACCCA 4620
          |||||||||||||||
Db 4292 --GGTCCCAGCAGGACTCCAAATTTAGGCAGCAGGCA 4328

Qy 4621 TATGGGATGGGATATAAAGGGCTGAGCACTGAGAGGGCTGAGCACTGAGAGGGAGATAT 4680
          |||||||||||||||||||||||||||||
Db 4329 TATGGGATGGGATATAAAGGGCTGAGCACTGAGAGGAT 4388

Qy 4681 GGTAAGAGGGAGTTTCGGGGTTCGGGGTTCGGGGTGGGGGGTTCAGGCGGTCCACCACCAGAA 4740
          |||||||||||||||
Db 4389 GGTAAGAGGGAGTTTCGGGGTTCGGGGTGGGGGGTTCAGGCGGTCCACCACCAGAA 4448

Qy 4741 GGAAACTGCCTTTCCCTGGAAGTGGGGTTCAGGCCGGTCGACAGGGTGGCCTT 4800
          ||||||||||||||
Db 4449 GGAAACTGCCTTTCCCTGGAAGTGGGGTTCAGGCCGGTCGACAGGGTGGCCTT 4508

Qy 4801 CCACCAGCCTGGAGTTCTCAGTGCGCAGGAGGTTTCCAAGAGAACACTGATGATGCCCCT 4860
          ||||||||||||||
Db 4509 CCACCAGCCTGGAGTTCTCAGTGCGCAGGAGGTTTCCACAGAGAACACTGATGATGCCCCT 4568

Qy 4861 TCCCTTACGGCTGTCTTCTCCATCTTCGCGTGTCCTTCTTCCTTCCTTGGTTTATC 4920
          ||||||||||||||
Db 4569 TCCCTTACGGCTGTCTTCTCCATCTTCGCGTGTCCTTCTTCCTTCCTTGGTTTATC 4628

Qy 4921 TTGGCTCTTCGTCTCTCAGCAAGATTTGCCCGTTCCTTCCTTCCTTCCTTCTCTACTGT 4980
          ||||||||||||||
Db 4629 TTGGCTCTTCGTCTCTCAGCAAGATTTGCCCGTTCCTTCCTTCCTTCCTTCTCTACTGT 4688

Qy 4981 CTCCGTGCCTTGCCCTTGCCTTTGCCTTGCGTGTCCTTCTTCCTTCCTTCCTTCACC 5040
          ||||||||||||||
Db 4689 CTCCGTGCCTTGCCCTTGCCTTTGCCTTGCGTGTCCTTCTTCCTTCCTTCCTTCACC 4748

Qy 5041 TTTTCTTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100
          ||||||||||||||
Db 4749 TTTTCTTCCCCTTCATTGTATTCATCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4808

Qy 5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5160
          ||||||||||||||
Db 4809 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4868

Qy 5161 CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCCACCCTTATGTAAACAATCTTCCA 5220

Qy 3001 TGAAAAGACAGCAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTGGGGTGCTTGATGGTG 3060
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3001 TGAAAAGACAGCAGCCCTGGAGGACAGGGGTTGTCTCTGAGCCTTGGGGTGCTTGATGGTG 3060

Qy 3061 CCACAAAGGAGGGCATGAGTGTGAGTATAAGGCCCCAGGAGCGTTAGAGAAGGGCACTTG 3120
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 3061 CCACAAAGGAGGGCATGAGTGTGAGTATAAGGCCCCAGGAGCGGTTAGAGAGAAGGGCACTTG 3120

Qy 3121 GGAAGGGGTCAGTTGCGAGAGCCCTATCCATCGGAATCTCATGGGGCCTGGGGCCAACTGGTG 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db 3121 GGAAGGGGTCAGTTGCGAGAGCCCTATCCATCGGAATCTCATGGGGCCTGGGGCCAACTGGTG 3180

Qy 3181 TAAATCTCTGGGCCTGCCCAGGCATTCAAAGCACACTGCCATCTCTGCCAGCCTGGGGA 3240
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3181 TAAATCTCTGGGCCTGCCCAGGCATTCAAAGCACACTGCCATCTCTGCCAGCCTGGGGA 3240

Qy 3241 GGGGGAAGGGAGCAACCCCCCACTTATACCCCTTTCTCCCTCAGCCCCAGGATTAACACCT 3300
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3241 GGGGGAAGGGAGCAACCCCCCACTTATACCCCTTTCTCCCTCAGCCCCAGGATTAACACCT 3300

Qy 3301 CTGGCCTTCCCCCTTCCCACCTCCCATCAGGAGTGAGAGGTTGCAGAGGGAGGGTAAAAA 3360
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3301 CTGGCCTTCCCCCTTCCCACCTCCCATCAGGAGTGAGAGGTTGCAGAGGGAGGGTAAAAA 3360

Qy 3361 CCTACATGTCCAAACATCATGGTGCACGATATATGGATCAGTAGTGTAGAGGCAAGAAA 3420
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3361 CCTACATGTCCAAACATCATGGTGCACGATATATGGATCAGTAGTGTAGAGGCAAGAAA 3420

Qy 3421 GGAAATCTGCAGCTTAACTGGGTTAATGTGTAAAGTCGTGTCGTGTGTGTCGTGTGTGTCTG 3480
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3421 GGAAATCTGCAGCTTAACTGGGTTAATGTGTAAAGTCGTGTCGTGTGTGTCGTGTGTGTCTG 3480

Qy 3481 ACTGAAAACGGGCATGGCTGTGCAGCTGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCA 3540
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3481 ACTGAAAACGGGCATGGCTGTGCAGCTGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCA 3540

Qy 3541 GGTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3541 GGTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT 3600

Qy 3601 TGGATGATGGCCTCGATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT 3660
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3601 TGGATGATGGCCTCGATCTCAAGGACCATGGAAAATAGAATGGACACTCTATATGTGTCT 3660

Qy 3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGACACCTGTCTAGAGATGTGGCAACAGAGAC 3720
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3661 CTAAGCTAAGGTAGCAAGGTCTTTGGAGACACCTGTCTAGAGATGTGGCAACAGAGAC 3720

Qy 3721 TACAGACAGTATGTGTACAGAGTAAGGAGGGGGTGTAGAATTCTTCTTACTA 3780
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3721 TACAGACAGTATGTGTACAGAGTAAGGAGGGGGTGTAGAATTCTTCTTACTA 3780

Qy 3781 TCAAAGGGAAACTGAGTCGTGCACCTGGAAGTCCTGCAAAGTCATCATGACTT 3840
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3781 TCAAAGGGAAACTGAGTCGTGCACCTGGAAGTCCTGCAAAGTCATCATGACTT 3840

Qy 3841 TGTCTCTGGGAGCAGTGTACAGAGTAAGGGAAGGGGAGAGGCTCCCTCA 3900
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3841 TGTCTCTGGGAGCAGTGTACAGAGTAAGGGAAGGGGAGAGGCTCCCTCA 3900

Qy 3901 GCCTGAAGCTATGCAGATAGCCAGCGGGTTGAAAGGGGAAGGGGAGAGCCTGGGATGGGAGC 3960
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3901 GCCTGAAGCTATGCAGATAGCCAGCGGGTTGAAAGGGGAAGGGGAGAGCCTGGGATGGGAGC 3960

Qy 3961 TTGTGTGTTGGAGGCAGGGGACAGATATTAAGCCTGGAAGAGAAGGTGACCCTTACCCAG 4020
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 3961 TTGTGTGTTGGAGGCAGGGGACAGATATTAAGCCTGGAAGAGAAGGTGACCCTTACCCAG 4020

Qy 4021 TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTGGGTAGGGGGAGGTG 4080
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db 4021 TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTGGGTAGGGGGAGGTG 4080

```
Db    4869 CTGTGTCAGAGTGCTGAGAATCACACCTGGGGTTCCACCCTTATGTAATCATCTTCCA 4928
           |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Qy    5221 GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCCTCCTTACCTTCCTCCACCCCTGGCTTGTC 5280
           ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db    4929 GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCCTCCTTACCTTCCTCCACCCCTGGCTTGTC 4988

Qy    5281 CTGTTCCATCCTGTCAGGATCTTCTAGATTGGTCTCCCAGCCTCTGTACTCCTCTTCCT 5340
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4989 CTGTTCCATCCTGTCAGGATCTTCTAGATTGGTCTCCCAGCCTCTGTACTCCTCTTCCT 5048

Qy    5341 GCCTGTTCCTCCTCTCGTCCAGCTGCGCCACTGTGGTGCCTGTTCCAGCTGTGGTCCAC 5400
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5049 GCCTGTTCCTCCTCTCGTCCAGCTGCGCCACTGTGGTGCCTGTTCCAGCTGTGGTCCAC 5108

Qy    5401 ATTCTTCAGGATTCTTGAAAAGTTAACCAGGTTGAGAATGTTTCCCCTGTAGACAGCAGA 5460
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5109 ATTCTTCAGGATTCTTGAAAAGTTAACCAGGTTGAGAATGTTTCCCCTGTAGACAGCAGA 5168

Qy    5461 TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCTCTCTTTCTGCCCAGCTGCCCCGGCA 5520
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5169 TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCTCTCTTTCTGCCCAGCTGCCCCGGCA 5228

Qy    5521 CTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTCTGACAGAGAAGCAGGCACT 5580
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5229 CTCTTAGCAAACCTCAGGCACCCTTACCCCACATAGACCTCTGACAGAGAAGCAGGCACT 5288

Qy    5581 TTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGAGAGGCAGGGAAGTGGT 5640
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5289 TTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGGTGTAAAAGAGAGGCAGGGAAGTGGT 5348

Qy    5641 GGTGTAGGAGAAGTCAGGACTTCACATAGAAGCTAGCCCACCACCAGAAATGACAGACAGA 5700
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5349 GGTGTAGGAGAAGTCAGGACTTCACATAGAAGCTAGCCCACCACCAGAAATGACAGACAGA 5408

Qy    5701 TCCCTCCTATCTCCCCATAAAGAGTTTGAGTCGAC 5735
           |||||||||||||||||||||||||||||||||||
Db    5409 TCCCTCCTATCTCCCCATAAAGAGTTTGAGTCGAC 5443
```

RESULT 4

US-10-944-375-15

; Sequence 15, Application US/10944375
; Publication No. US20050066381A1
; GENERAL INFORMATION:
; APPLICANT: Cincinnati Children's Hospital Medical Center
; APPLICANT: University of Cincinnati
; APPLICANT: Molkentin, Jeffery D
; APPLICANT: Kranias, Evangelia G
; TITLE OF INVENTION: REGULATION OF CARDIAC CONTRACTILITY AND HEART FAILURE PROPENSITY
; FILE REFERENCE: 9761M#L
; CURRENT APPLICATION NUMBER: US/10/944,375
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: US 60/503,853
; PRIOR FILING DATE: 2003-09-19
; SOFTWARE: PatentIn version 3.3
; NUMBER OF SEQ ID NOS: 15
; SEQ ID NO 15
; LENGTH: 5443
; TYPE: DNA
; ORGANISM: Mus musculus

US-10-944-375-15

Query Match 88.0%; Score 5045; DB 19; Length 5443;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

```
Qy       1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCCACCAGGTGCCCTAGTCTCAATTCAGT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGATCCTGCAAGGTCACACAAGGGTCTCCACCCACCAGGTGCCCTAGTCTCAATTCAGT 60

Qy      61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTTGGACTTTGTTTTTAT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTCCATGCCTTGTTCTCACAATGCTGGCCTCCCCAGAGCTAATTTGGACTTTGTTTTTAT 120
```

```
Qy     121 TTCAAAAGGGCCTGAATGAGGAGGTAGATCTTGTGTCACCCAGCTCAAGGGTGCCGTGA 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TTCAAAAGGGCCTGAATGAGGAGGTAGATCTTGTGTCACCCAGCTCAAGGGTGCCGTGA 180

Qy     181 AGCCCTCAGACCTGGAGCCCTTAGGTGGAAGCAGAATAAAGCAATTT 240
           |||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGCCCTCAGACCTGGAGCCCTTAGGTGGAAGCAGAATAAAGCAATTT 240

Qy     241 TCCTTAAAGCCAAAATCCTGCCTCTTCTCTGCCTCTTCTGACTCTTCT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TCCTTAAAGCCAAAATCCTGCCTCTTCTCTGCCTCTTCTGACTCTTCT 300

Qy     301 AGGGTGGGGAGGTGGGGCTTGGAAGAGAAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 AGGGTGGGGAGGTGGGGCTTGGAAGAGAAGGTGGGGAAGTGGCAAAAGCCGATCCCTAG 360

Qy     361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCTCCAGCC 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GGCCCTGTGAAGTTCGGAGCCTTCCCTGTACAGCACTGGCTCATAGATCCTCCTCCAGCC 420

Qy     421 AAACATAGCAAGAAGTGATACCTTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 AAACATAGCAAGAAGTGATACCTTCCTTTGTGACTTCCCAGGCCCAGTACCTGTCAGGTT 480

Qy     481 GAAACAGGATTTAGAGAAGCCTCGAACTCACCTGAGCTCGAGCTCATCATCCACCAAGCA 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GAAACAGGATTTAGAGAAGCCTCGAACTCACCTGAGCTCGAGCTCATCATCCACCAAGCA 540

Qy     541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAATGAGAGCTGGGCCAC 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 AGCACCTAGGTGCCACTGCTAGTTAGTATCCTGATAATATGCAGAATGAGAGCTGGGCCAC 600

Qy     601 AGAAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCGGCAAAGTATGACCCCC 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AGAAGTCCTGGGGTGTAGGAACTGACCAGTGACTTTTCAGTCGGCAAAGTATGACCCCC 660

Qy     661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCCATCCCAGATCCCATCCCTAAGAGGACATG 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 TCAGCAGATGTAGTAATGTCCCCTTAGATCCCATCCCAGATCCCATCCCTAAGAGGACATG 720

Qy     721 GGATGAGAGATGTAGTCATGTGCATTCCAAACACAGCTATCTTCTTCCTCCCCATCCTCCAG 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GGATGAGAGATGTAGTCATGTGCATTCCAAACACAGCTATCTTCTTCCTCCCCATCCTCCAG 780

Qy     781 TTCCACTTAGCCAGGAGACAGTAACCTTAGCCTTAGCCTTAGCCTTTAGGAATGTGGGTC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TTCCACTTAGCCAGGAGACAGTAACCTTAGCCTTAGCCTTAGCCTTTAGGAATGTGGGTC 840

Qy     841 GACACACCCCTGGTCTGCAGTATTCATTTCATTTCCTTCACGTCGACTTCCAT 900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GACACACCCCTGGTCTGCAGTATTCATTTCATTTCCTTCACGTCGACTTCCAT 900

Qy     901 TTGCAAGGCTTTTGACCCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 960
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTGCAAGGCTTTTGACCCTGCTGGAAGATAGAGTTTGGCCCTAGGTGTGGCAA 960

Qy     961 GCCATCTCAAGAGAAAGCAGACAACAGGGGACCAGATTTGGAAGGATCAGGAACTAAA 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GCCATCTCAAGAGAAAGCAGACAACAGGGGACCAGATTTGGAAGGATCAGGAACTAAA 1020

Qy    1021 TCACTGGGGGGCCTGGGGTGGTAGAAAAAGAGTGAGTGAGTGTCCGTTCCAGCTAAGCCAAGC 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TCACTGGGGGGCCTGGGGTGGTAGAAAAAGAGTGAGTGAGTGTCCGTTCCAGCTAAGCCAAGC 1080

Qy    1081 TAGTCCCGAGATACTCTGCCACCAGGTCTTAGGAATGTGGGTC 1140
           |||||||||||||||||||||||||||||||||||||||||||
Db    1081 TAGTCCCGAGATACTCTGCCACCAGGTCTTAGGAATGTGGGTC 1140

Qy    1141 TGAAAGACAAATGGGATTGGAAGACATCTCTTTGAAGTCTCTTTGAGTCTCCCCTCAACCCCACTACAGAC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 TGAAAGACAAATGGGATTGGAAGACATCTCTTTGAAGTCTCTTTGAGTCTCCCCTCAACCCCACTACAGAC 1200
```

```
Qy  1201  ACACTCGTGTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACCACTGAGCTAG  1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1201  ACACTCGTGTGTGGCCAGACTCCTGTTCAACAGCCCTCTGTGTTCTGACCACTGAGCTAG  1260

Qy  1261  GCAACCAGAGCATGGGCCCTGTCTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG  1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1261  GCAACCAGAGCATGGGCCCTGTCTGCTGAGGATGAAGAGTTGGTTACCAATAGCAAAAACAG  1320

Qy  1321  CAGGGGAGGGGAGAACGAGAGAGATAAGGAAGGAAGGAAAGGCCAGTCAGTCAATCAGA  1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1321  CAGGGGAGGGGAGAACGAGAGAGATAAGGAAGGAAGGAAAGGCCAGTCAGTCAATCAGA  1380

Qy  1381  TGCAGTCAGAAGAGATGGGAAGCCAACACAGCTTGAGCAGAGAAACAGAAAAAGGGAG  1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1381  TGCAGTCAGAAGAGATGGGAAGCCAACACAGCTTGAGCAGAGAAACAGAAAAAGGGAG  1440

Qy  1441  AGATTCTGGGCATAAGAGGGCCACAGAAGAAGAGCCCAGGCCCCCAAGTCTCCTCTTT  1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1441  AGATTCTGGGCATAAGAGGGCCACAGAAGAAGAGCCCAGGCCCCCAAGTCTCCTCTTT  1500

Qy  1501  ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC  1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1501  ATACCCTCATCCCGTCTCCCAATTAAGCCCACTCTTCTTCCTAGATCAGACCTGAGCTGC  1560

Qy  1561  AGCGAAGAGACCCGTAGGGAGGATCACACTGGATGAGAGATGTGTGAGAGTCCAGG  1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1561  AGCGAAGAGACCCGTAGGGAGGATCACACTGGATGAGAGATGTGTGAGAGTCCAGG  1620

Qy  1621  GCAACCTAAGAGCCAGAGCCTAAAAGCAAGAGCAAGATAAAGTGCTTCAAAGGTGGCCAGG  1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1621  GCAACCTAAGAGCCAGAGCCTAAAAGCAAGAGCAAGATAAAGTGCTTCAAAGGTGGCCAGG  1680

Qy  1681  CTGTGCACACAGAGGGTCGAGGACTGGTGGTAGAGCCTCAAGATAAGGATGATGCTCAGA  1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1681  CTGTGCACACAGAGGGTCGAGGACTGGTGGTAGAGCCTCAAGATAAGGATGATGCTCAGA  1740

Qy  1741  ATGGGCGGGGGGGGGATTCTGGGGGGGGGGATTCTGGGGGGAGAAGGTGAGAAGGAGCCTGGAACAG  1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1741  ATGGGCGGGGGGGGGATTCTGGGGGGGGGGATTCTGGGGGGAGAAGGTGAGAAGGAGCCTGGAACAG  1800

Qy  1801  AGAATCTGGAAGCGCTGGAAACGATACCATAAAGGGAGAACCCAGGCTACCTTTAGATG  1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1801  AGAATCTGGAAGCGCTGGAAACGATACCATAAAGGGAGAACCCAGGCTACCTTTAGATG  1860

Qy  1861  TAAATCATGAAAGACAGGAGACAGGAGACAGAGAGGGAGCTGAGAGACTAGAAGGACCCCGGGGCAAGA  1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1861  TAAATCATGAAAGACAGGAGACAGGAGACAGAGAGGGAGCTGAGAGACTAGAAGGACCCCGGGGCAAGA  1920

Qy  1921  CATGGAAGCCAAGGACAAGCAACCCAGGTTGAGCGCGTCCGTGAAATCAGCGCTGCTGAAGGCGAGAG  1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1921  CATGGAAGCCAAGGACAAGCAACCCAGGTTGAGCGCGTCCGTGAAATCAGCGCTGCTGAAGGCGAGAG  1980

Qy  1981  CCCTGGTATGAGCACCAGGAACAGCAGGAACAGGAACAGAGAAG  2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1981  CCCTGGTATGAGCACCAGGAACAGCAGGAACAGGAACAGAGAAG  2040

Qy  2041  GTAGACACAGGAACAGAGACAGAGACAGGCGGGGGACGGCGGGGGACGGCGGGGGACCTAGGA  2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2041  GTAGACACAGGAACAGAGACAGAGACAGGCGGGGGACGGCGGGGGACGGCGGGGGACCTAGGA  2100

Qy  2101  CTGTGGCCAGAGCGTTCCATCTGTGTCCAGATTCATCATGTTCATCAGACTGTCATCAGGGC  2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2101  CTGTGGCCAGAGCGTTCCATCTGTGTCCAGATTCATCATGTTCATCAGACTGTCATCAGGGC  2160

Qy  2161  TGGCTTGGGAGGTCAGTCAGCTGGAAAGAGTATGTGAGAGCAAGGGGACAAGGGGACAAGGGGCCTAGGA  2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2161  TGGCTTGGGAGGTCAGTCAGCTGGAAAGAGTATGTGAGAGCAAGGGGACAAGGGGACAAGGGGCCTAGGA  2220

Qy  2221  AAGGAAGAGAGGGCCACAACCAGGCCAACAGGGCCAACAGGGCCACAGGGCCACACAAGAGGAGGGACTGAGTTAACTC  2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2221  AAGGAAGAGAGGGCCACAACCAGGCCAACAGGGCCAACAGGGCCACAGGGCCACACAAGAGGAGGGACTGAGTTAACTC  2280
```

```
Qy  2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGCTGGACCACCATCCCCATGA  2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2281  CTTCCTTGTTGCATCTTCCATAGGAGGCAGTGGGAACTCTGTGCTGGACCACCATCCCCATGA  2340

Qy  2341  GCCCCCACTACCCATACCAAGTTTGGCTTGGTGCATTCTAGGTTCCCTGAGGACAGAG  2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2341  GCCCCCACTACCCATACCAAGTTTGGCCTGAGTTTGGCCATTCTAGGTTCCCTGAGGACAGAG  2400

Qy  2401  CCTGGCCTTGTCTCTTGACCTGACCCAAGCTGACCCAATGTTCTCAGTAGTCTTTGAAT  2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2401  CCTGGCCTTGTCTCTTGACCTGACCCAAGCTGACCCAATGTTCTCAGTACCTTATCAT  2460

Qy  2461  GCCCTCAAGAGCTTGAGACACCAGGCAGTGACATATTAGGCTAACCCTGGAGCT  2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2461  GCCCTCAAGAGCTTGAGAACCAGGCAGTGACATATTAGGCTAACCCTGGAGCT  2520

Qy  2521  TGCACACAGAGCCTCAAGTGACCTGCAGGGACACAGGTGAATCAGGTC  2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2521  TGCACACAGAGCCTCAAGTGACCTCAGGGACACAGGTCAGGTC  2580

Qy  2581  CCAAAGAGCAACCATTTGCATAGGTGGCTGCAAATGGGAAATGCAAGTTGAATCAGGTC  2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2581  CCAAAGAGCAACCATTTGCATAGGTGGCTGCAAATGGGAAATGCAAGTTGAATCAGGTC  2640

Qy  2641  CCTTCAAGAATACTGCATGCAAGAACCCCTGGAGAGGGGGTATGCTCCTGCCC  2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2641  CCTTCAAGAATACTGCATGCAAGAACCTAGACCCCTGAGAGAGGGTATGCTCCTGCCC  2700

Qy  2701  CCACCCCACCATAAGGGAGGTGAACTATCCTAGGGGGCTGGCGACCTTGGGGAGACACCAC  2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2701  CCACCCCACCATAAGGGAGGTGAACTATCCTAGGGGGCTGGCGACCTTGGGGAGACACCAC  2760

Qy  2761  ATTACTGAGAGTGCTGAGCCTGAGAAAAACTGACCGCCCTGTGTCCTGCCCACCTCCACAC  2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2761  ATTACTGAGAGTGCTGAGCCTGAGAAAAACTGACCGCCCTGTGTCCTGCCCACCTCCACAC  2820

Qy  2821  TCTAGAGCTATATTGAGAGTGACAGTAGGGTGGGAGCTGGAGCTGCAGGAGAGTGT  2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2821  TCTAGAGCTATATTGAGAGTGACAGTAGGGTGGGAGCTGGAGCTGCAGGAGAGTGT  2880

Qy  2881  TCCTGGGTGTGAGGGTGTAGGGTGGGGAAAGCCAGAGCTCTGGGAGTCTGCTTTGTCTCCTGAA  2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2881  TCCTGGGTGTGAGGGTGTAGGGTGGGGAAAGCCAGAGCTCTGGGAGTCTGCTTTGTCTCCTGAA  2940

Qy  2941  CACAATGTCACTTAGTTATAAGAGCCATGGCCTAAAGACCCTGCTAAAGACCTACGACCTC  3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2941  CACAATGTCACTTAGTTATAAGAGCCATGGCCTAAAGACCCTGCTAAAGACCTACGACCTC  3000

Qy  3001  TGAAAAGACAGCAGCGTTGTGAGCCGCTCCGGTTGTCTCTGAGCCTTGGGTGCTTGATGGTG  3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3001  TGAAAAGACAGCAGCGTTGTGAGCCGCTCCGGTTGTCTCTGAGCCTTGGGTGCTTGATGGTG  3060

Qy  3061  CCACAAAGGAGGGCATGAGGGCATGAGTTGAGTATAAGGCCCCAGAGCGTTAGAGAGGGGCACTTG  3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3061  CCACAAAGGAGGGCATGAGGGCATGAGTTGAGTATAAGGCCCCAGAGCGTTAGAGAGGGGCACTTG  3120

Qy  3121  GGAAGGGGTCAGTCTGCTGCAGAGCCCATGGAATCTGGAATCTGCTAAAGACCTCCCCTTCTCCCTCTACCTGGTG  3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3121  GGAAGGGGTCAGTCTGCTGCAGAGCCCATGGAATCTGGAATCTGCTAAAGACCTCCCCTTCTCCCTCTACCTGGTG  3180

Qy  3181  TAAATCTCTGGGCCTGCCAGGCATTCAAAGTTCATCGACTGTCAGGGC  3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3181  TAAATCTCTGGGCCTGCCAGGCATTCAAAGCTTATCGACTGTCAGGGC  3240

Qy  3241  GGCGGAAGGGAGCAACCCCCCACTTATACCCTTTCTCCCTCAGGATTAACACCTGGTG  3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3241  GGCGGAAGGGAGCAACCCCCCACTTATACCCTTTCTCCCTCAGGATTAACACCTGGTG  3300

Qy  3301  CTGGCCTTCCCCCTTCCCACCTCCCATCAGGAGTGCGAGGTTGCAGAGGGAGGGTAAAAA  3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3301  CTGGCCTTCCCCCTTCCCACCTCCCATCAGGAGTGCGAGGTTGCAGAGGGAGGGTAAAAA  3360

Qy  3361  CCTACATGTCCAAACATCATGGTGCACGATATATGGATCAGTAGTGTGTGAGAGGCAAGAAA  3420
```

```
Db  3361  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      CCTACACATGTCCAAACACATCATGTGTGCACGATATATGGATCAGTATGTGTGAGAGGCAAGAAA  3420

Qy  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTGTGAAAGTCTGTGTGCATGTGTGTGTGTCTG  3480
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3421  GGAAATCTGCAGGCTTAACTGGGTTAATGTGTGTGAAAGTCTGTGTGCATGTGTGTGTGTCTG  3480

Qy  3481  ACTGAAAACGGGCATGGCTGTGCAGCTGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCA  3540
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3481  ACTGAAAACGGGCATGGCTGTGCAGCTGTTCAGTTCTGTGCGTGAGGTTACCAGACTGCA  3540

Qy  3541  GGTTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT  3600
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3541  GGTTTGTGTGTAAATTGCCCAAGGCAAAGTGGGTGAATCCCTTCCATGGTTTAAAGAGAT  3600

Qy  3601  TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAACTCTATATGTGTCT  3660
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3601  TGGATGATGGCCTGCATCTCAAGGACCATGGAAAATAGAACTCTATATGTGTCT  3660

Qy  3661  CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC  3720
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3661  CTAAGCTAAGGTAGCAAGGTCTTTGGAGGACACCTGTCTAGAGATGTGGGCAACAGAGAC  3720

Qy  3721  TACAGACAGTATCTGTACAGGAGAGGGGGTGTAGAATTCTCTTACTA  3780
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3721  TACAGACAGTATCTGTACAGGAGAGGGGGTGTAGAATTCTCTTACTA  3780

Qy  3781  TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGGCTCTCCCTAGACATCATGACTT  3840
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3781  TCAAAGGGAAACTGAGTCGTGCACCTGCAAAGTGGATGGCTCTCCCTAGACATCATGACTT  3840

Qy  3841  TGTCTCTTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGATGTAGGAGGCTCCCCTCA  3900
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3841  TGTCTCTTGGGGAGCCAGCACTGTGGAACTTCAGGTCTGAGAGATGTAGGAGGCTCCCCTCA  3900

Qy  3901  GCCTGAAGCTATGCAGCAGGGTTGAAAGGGGTTGAAAGGGTTGAAAGGTGAAAGGAGAGCCTGGAGC  3960
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3901  GCCTGAAGCTATGCAGCAGGGTTGAAAGGGGTTGAAAGGGTTGAAAGGTGAAAGGAGAGCCTGGAGC  3960

Qy  3961  TTGTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAAGTGAAGAGAGGTGCACCCTTACCCAG  4020
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3961  TTGTGTGTTGGAGGCAGGGACAGATATTAAGCCTGGAAGTGAAGAGAGGTGCACCCTTACCCAG  4020

Qy  4021  TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTGGGTAGGGGAGGTG  4080
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4021  TTGTTCAACTCACCCTTCAGATTAAAAATAACTGAGGTAAGGGCCTGGGTAGGGGAGGTG  4080

Qy  4081  GTGTGAGACGCTCCTGTCTCCTGTCCATGCCCTGAGGCCCTTGGGGGAGGAGGAATGT  4140
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4081  GTGTGAGACGCTCCTGTCTCCTGTCCATCTCGCCCATCGGCCCTTGGGGGAGGAGGAATGT  4140

Qy  4141  GCCCAAGGACTAAAAAAGGCCATGGAGCCAGAGGGCCAACAGACCTTCATGG  4200
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4141  GCCCAAGGACTAAAAAAGGCCATGGAGCCAGAGGGCCAACAGACCTTCATGG  4200

Qy  4201  GCAAACCTGGGCCGTAGTGATCGATTGACAAGAACTCGGCCAATGCAAGGGATCAAAGGAGGA  4260
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4201  GCAAACCTGGGCCGTAGTGATCGATTGACAAGAACTCGGCCAATGCAAGGGATCAAAGGAGGA  4260

Qy  4261  TTCTAACGGACAGGAGGGAAGTGGGAGGGAG---------------------  4320
      |||||||||
Db  4261  GGAGCCAGGACAGGAGGGAAGTGGGAGGGAG  4291

Qy  4321  AAGTGAGGTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGCTTACCACTC  4380
      -----------------------------------------
Db  4292                                                         4291

Qy  4381  CCTATCAGTGATAGAGAAAGTGAGGTTACCACTCCTATCCATCAGTGATGATAGAGAA  4440
      --------------------------------------------------
Db  4292                                                         4291

Qy  4441  AAGTGAAAGTCGAGTTACCACTCCCTATCCATCAGTGATGATAGAGAAAGTGAAAGTCGAGTTA  4500
```

```
Db  4292  ---------------------------------------------------------  4291

Qy  4501  CCACTCCCTATCCAGTCGATGATAGAGAAAGTGAGGTTACCACTCCTATCCATCAGTGAT  4560

Db  4292  ---------------------------------------------------------  4291

Qy  4561  AGAGAAAGTGAAAGTCGAGCTCCAGCAGAGACTCCAAATTTAGGCAGCAGGCA  4620
              |||  ||||||||||||||||||||||||||||||||||
Db  4328        GGTCCCAGCAGAGAGACTCCAAATTTAGGCAGCAGGCA  4328

Qy  4621  TATGGGATGGATATAAAGGGGCTGGAGCACTGAGAGCTGTCAGAGATTTCTCAACCCA  4680
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4329  TATGGGATGGATATAAAGGGGCTGGAGCACTGAGAGCTGTCAGAGATTTCTCCAACCCA  4388

Qy  4681  GGTAAGAGGGAGTTCGGGTGGGTGGGGGCTCTTCACCACCAGACCTCTCCCACCACTAGAA  4740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4389  GGTAAGAGGGAGTTCGGGTGGGTGGGGGCTCTTCACCCACCAGACCTCTCCCACCACTAGAA  4448

Qy  4741  GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGGGTTGCCTT  4800
              |||||||||||||||||||||||||||||||||||||||||||||||||
Db  4449  GGAAACTGCCTTTCCTGGAAGTGGGGTTCAGGCCGGTCAGGGTTGCCTT  4508

Qy  4801  CCACCAGCTGGGAAGTTCTCAGTGGCAGGGAGTTTCCACAAGAAACACTGGATGCCCCT  4860
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4509  CCACCAGCTGGGAAGTTCTCAGTGGCAGGGAGTTTCCACAAGAAACACTGGATGCCCCT  4568

Qy  4861  TCCCTTACGCTGTCTTCTCCATCTTCCTCCTGGGGATGCTCCTCCCCGTTCGTTGGTTTATC  4920
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4569  TCCCTTACGCTGTCTTCTCCATCTTCCTCCTGGGGATGCTCCTCCCCGTTCGTTGGTTTATC  4628

Qy  4921  TTGGCTCTTCGTCTTCAGCAAGATTGCCCTGGCGTGTCGTCCTCCTTCCTCCACTCTGT  4980
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4629  TTGGCTCTTCGTCTTCAGCAAGATTGCCCTGGCGTGTCGTCCTCCTTCCTCCACTCTGT  4688

Qy  4981  CTCCGTGCCTTGCCTTCTTGGCGTGTCCTTTCCTTTCCTTTCCTTTCTCACTCACC  5040
              ||||||||||||||||||||||||||||||||||||||||||||
Db  4689  CCACCGTGCCTTGCCTTCTTGGCGTGTCCTTTCCTTTCCTTTCTCACTCACC  4748

Qy  5041  TTTTCTCCCCTTCCTCATTTGTATTCATCCTCCCTTCCTTCCTTCCTCCTTC  5100
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4749  TTTTCTCCCCTTCCTCATTTGTATTCATCCTCCCTTCCTTCCTTCCTCCTTC  4808

Qy  5101  CTTCCTTCCTCCTTCCTTCCTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  5160
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4809  CTTCCTTCCTCCTTCCTTCCTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC  4868

Qy  5161  CTGTGTCAGAGTGCTGAGAATCACACACTCGGGGTCCCACCCTTATGTAAACAATCTTCCA  5220
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4869  CTGTGTCAGAGTGCTGAGAATCACACACTCGGGGTCCCACCCTTATGTAAACAATCTTCCA  4928

Qy  5221  GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTCTCTACCTTCCTCACCCCTGGCTTGTC  5280
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4929  GTGAGCCACAGCTTCAGTGCTGCTGGGTGCTCTCTCTACCTTCCTCACCCCTGGCTTGTC  4988

Qy  5281  CTGTTCCATCCTGTCAGGATCTCTAGATTGGTCTCCAGCCTCTCTTCCTTCCT  5340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  4989  CTGTTCCATCCTGTCAGGATCTCTAGATTGGTCTCCAGCCTCTCTTCCTTCCT  5048

Qy  5341  GCCTGTTCCTCCTCCTCTGCCAGCTGCGCCACTGTGGTGCCACTGTGTCCAC  5400
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5049  GCCTGTTCCTCCTCTCTGTCGTGCCAGCTGCGCCGCCACTGTGGTGCCACTGTGTCCAC  5108

Qy  5401  ATTCTTCAGGATTCTCTGGAAAGTGAAAGTTCCCCTGAAAGTGTTCCCCTCAGACAGA  5460
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5109  ATTCTTCAGGATTCTCTGGAAAGTGAAAGTTAACCAGTGAGAATGTTCCCCTGCCCAGCAGA  5168

Qy  5461  TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTTCTGCCCAGGCCA  5520
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5169  TCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTTCTGCCCAGGCCA  5228

Qy  5521  CTCTTAGCAAACCTCAGGCACCCTTACCCCACCCTTACACCTCTGACAGACCTGCGAAGCACT  5580
              ||||||||||||||||||||||||||||||||||||||||||||||||||
Db  5229  CTCTTAGCAAACCTCAGGCACCCTTACCCCACCCTTACAGACCTCTGACAGACCTGCGCCGGCACT  5288
```

```
Qy   4247  CGATACCCTTCTTCTTCTTCTTCTCAACGACGAGGAGGAACTCGAGTTTACCACTCCTATCAGTG  4306
                  ||||||||||||||||
Db    481  GGATCAAAGGAGGAGGAGCCAGGACAGGAGGGAGGGAAGTGGGAGGGAG-------------------  525

Qy   4307  ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGT  4366
Db    526  -------------------------------------------------------------  525

Qy   4367  CGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTA  4426
Db    526  -------------------------------------------------------------  525

Qy   4427  TCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGT  4486
Db    526  -------------------------------------------------------------  525

Qy   4487  GAAAGTCGAGTTTACCACTCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC  4546
Db    526  -------------------------------------------------------------  525

Qy   4547  TCCCTATCAGTGATAGAGAAAAGTGAGCTCGGTACCACGAGAGGACTCCAAAT  4606
                                   ||||||||||||||||||||
Db    526  ----------------------GGTCCCCAGGAGGACTCCAAAT  548

Qy   4607  TTAGGCAGCAGGCATATGGGATGGGATATAAAGGGCCTGGAGCACTGAGAGCTGTGCAGAG  4666
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    549  TTAGGCAGCAGGCATATGGGATGGGATATAAAGGGCCTGGAGCACTGAGAGCTGTGCAGAG  608

Qy   4667  ATTTCTCCAACCCAGGTAGAGAGGAGTTTCGGGTTGGGGGCTCTTCACCCACCACCAGACCT  4726
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    609  ATTTCTCCAACCCAGGTAAGAGAGGAGTTTCGGGTTGGGGGCTCTTCACCCACCACCAGACCT  668

Qy   4727  CTCCCCACCTAGAAGAAACTGCCTTCCTGGAAGTGGGTTCAGGCCGGTCAGGAGATCT  4786
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    669  CTCCCCACCTAGAAGAAACTGCCTTCCTGGAAGTGGGTTCAGGCCGGTCAGGAGATCT  728

Qy   4787  GACAGGGGTGGCCTTCCACAGCCTGGGAGTTCTCAGTGCGCAGGAGTTTCCACAAGAAA  4846
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    729  GACAGGGGTGGCCTTCCACAGCCTGGGAGTTCTCAGTGCGCAGGAGTTTCCACAAGAAA  788

Qy   4847  CACTGGATGCCCCTTCCCTTACGCTCGTCTTCCTCCTCCTGGGGATGCTCCTCCC  4906
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    789  CACTGGATGCCCCTTCCCTTACGCTCGTCTTCCTCCTCCTGGGGATGCTCCTCCC  848

Qy   4907  CGTCTTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTCGTCCACTCCA  4966
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    849  CGTCTTGGTTTATCTTGGCTCTTCGTCTTCAGCAAGATTTGCCCTGTCGTCCACTCCA  908

Qy   4967  TCTTTCTCTACTGTCTCCGTGCCTTCTTGCCGTGTCCTTCCTTTCCACCCA  5026
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    909  TCTTTCTCTACTGTCTCCGTGCCTTCTTGCCGTGTCCTTCCTTTCCACCCA  968

Qy   5027  TTTCTCACTTCACCTTTTCTCCCCCTTCATTGTATTCATCCTTCTTCCTTCTTCCTT  5086
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    969  TTTCTCACTTCACCTTTTCTCCCCCTTCATTGTATTCATCCTTCTTCCTTCTTCCTT  1028

Qy   5087  TCCTTCTTCCTTCTGTGTCAGAGTGCTGAGAATCACACCTGGGTTCCACCCTTATG  5146
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1029  TCCTTCTTCCTTCTGTGTCAGAGTGCTGAGAATCACACCTGGGTTCCACCCTTATG  1088

Qy   5147  TCCTTCCTTCCTTCTGGTCAGAAGTGCTGAGATCTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG  5206
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1089  TCCTTCCTTCCTTCTGGTCAGAAGTGCTGAGATCTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG  1148

Qy   5207  TAAACAATCTTCAGTGAGCCAACTTCAGTGCTCGGGGCGCGAGGGCAA  5266
           |||||||||||||||||||||||||||||||||||||||||||||
Db   1149  TAAACAATCTTCAGTGAGCCAACTTCAGTGCTCGGGGCGCGAGGGCAA  1208

Qy   5267  CCCCTGGCTGTCTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG  5326
           ||||||||||||||||||||||||||||||||||||||||||||||
Db   1209  CCCCTGGCTGTCTGTTCCATCCTGGTCAGGATCTCTAGATTGGTCTCCCAGCCTCTG  1268

Qy   5327  CTACTCCTTCCTTCCTGCCTGTTCCTCTCTGTCGTCGCCACTGTGTGCCACTGTGCCTTCGTTC  5386
```

```
Qy   5581  TTACATGGAGTCCTGGTGGGAGAGGCCATAGGCTACGGTGTAAAGAGGCAGGAAGTGGT  5640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5289  TTACATGGAGTCCTGGTGGGAGAGGCCATAGGCTACGGTGTAAAGAGGCAGGAAGTGGT  5348

Qy   5641  GGTGTAGGAAGTCAGGACTTCACATAGAAGCTAGCCCACCACCAGAAATGACAGACAGA  5700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5349  GGTGTAGGAAGTCAGGACTTCACATAGAAGCTAGCCCACCACCAGAAATGACAGACAGA  5408

Qy   5701  TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC  5735
           ||||||||||||||||||||||||||||||||||
Db   5409  TCCCTCCTATCTCCCCATAAGAGTTTGAGTCGAC  5443
```

RESULT 5
US-10-027-655-5
; Sequence 5, Application US/10027655
; Publication No. US20030204206A1
; GENERAL INFORMATION:
;   APPLICANT: Medtronic, Inc.
;   APPLICANT: Padua, Rodolfo
;   APPLICANT: Schu, Carl
;   APPLICANT: Bonner, Matthew
;   APPLICANT: Donovan, Maura
;   APPLICANT: Soykan, Orhan
;   TITLE OF INVENTION: Electrically Responsive Promoter System
;   FILE REFERENCE: P9406.00
;   CURRENT APPLICATION NUMBER: US/10/027,655
;   CURRENT FILING DATE: 2001-12-20
;   NUMBER OF SEQ ID NOS: 6
;   SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 1679
;   TYPE: DNA
;   ORGANISM: Mouse
US-10-027-655-5

Query Match   22.3%; Score 1279.8; DB 17; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;

```
Qy   3767  GAATTCTCTTACTATCCAAAGGAGAAACTGAGTCGTGCACCTGCAAAGTCGATGCTCTCCCT  3826
           |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  GAATTCTCTTACTATCCAAAGGAGAAACTGAGTCGTGCACCTGCAAAGTCGATGCTCTCCCT  60

Qy   3827  AGACATCATGACTTTGTCTCTGGGAGCCAGCACTGTGGAACTTCAGTCTGAGAGAGTA  3886
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  AGACATCATGACTTTGTCTCTGGGAGCCAGCACTGTGGAACTTCAGTCTGAGAGAGTA  120

Qy   3887  GGAGGCGCTCCCCCCAGCCTGAAGCTATGCGAAGATAGCCAGGGGTTGAAAAGGGGAGGAGAG  3946
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GGAGGCGCTCCCCCCAGCCTGAAGCTATGCGAAGATAGCCAGGGGTTGAAAAGGGGAGGAGAG  180

Qy   3947  CCTGGGATGGGAGCTTGTGTGTTGGAGGCAGGACAGATATTAAGCCTGGAAGAGAAAGG  4006
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  CCTGGGATGGGAGCTTGTGTGTTGGAGGCAGGACAGATATTAAGCCTGGAAGAGAAAGG  240

Qy   4007  TGACCCTTACCCAGTTGTTCAACTCACCCTTCACCTCACCTCAGATTAAAAATAACTGAGGTAAGGGCCT  4066
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241  TGACCCTTACCCAGTTGTTCAACTCACCCTTCACCTCACCTCAGATTAAAAATAACTGAGGTAAGGGCCT  300

Qy   4067  GGGTAGGGAGGTGGTGAGAGCGTGTGAGAAGCTCCTCGTGTCTCCTGTCTCCTGCATGCCCTGAGGCCCCTTTG  4126
           ||||||||||||||||||||||||||||||||||||||||  |||||||||||
Db    301  GGGTAGGGAGGTGGTGAGACGCTGTGAGACGCTCCTCGTGTCTCCTGTCTCCCATCGGCCCCTTTG  360

Qy   4127  GGGAGGAGGAATGCGCCCAAGACTAAAAAAGGCCATGGAGCCAGGGGGCGAGGGCAA  4186
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361  GGGAGGAGGAATGCGCCCAAGACTAAAAAAGGCCATGGAGCCAGGGGGCGAGGGCAA  420

Qy   4187  CAGACCTTTCATGGGCAAACCTTGGGCCCCGTAGTGAGTCGATCGATGACAAGAACTCGCCAAT  4246
           |||||||||||||||||||||||||||||||||||||||||||||||||  ||||||
Db    421  CAGACCTTTCATGGGCAAACCTTTGGGCCCCGTGTCCTTCCTGTCACCTCCAGAGCCAAG  480
```

```
Db   1269 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          CTACTCCTCTGCCTGTTCCTCTCTCTGTCTGCGCCACTGTGGTCGTGCCACTGTGGTCGTTC 1328

Qy   5387 CAGCTGTGTCCACCATTCTTCAGGATTCTTCAGGATTCTGAAAAGTTAACCAGGTGAGAATGTTCCC 5446
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1329 CAGCTGTGTCCACCATTCTTCAGGATTCTTCAGGATTCTGAAAAGTTAACCAGGTGAGAATGTTCCC 1388

Qy   5447 CTGTAGACAGCAGATCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTCTTCTCTGC 5506
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1389 CTGTAGACAGCAGATCACGATTCTCCCGGAAGTCAGGCTTCCAGCCCTCTCTTCTTCTCTGC 1448

Qy   5507 CCAGCTGCCCGGCCACTCTTAGCAAACCTCAGGCACCCTTACCCCCACCATAGACCTCTGACA 5566
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1449 CCAGCTGCCCGGCCACTCTTAGCAAACCTCAGGCACCCTTACCCCCACCATAGACCTCTGACA 1508

Qy   5567 GAGAAGCAGGCACTTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGTGTAAAAGA 5626
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1509 GAGAAGCAGGCACTTTACATGGAGTCCTGGTGGGAGAGCCATAGGCTACGTGTAAAAGA 1568

Qy   5627 GGCAGGGAAGTGTGGTGTGGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACCACCAG 5686
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1569 GGCAGGGAAGTGTGGTGTGGTGTAGGAAAGTCAGGACTTCACATAGAAGCCTAGCCCACCACCAG 1628

Qy   5687 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 5731
          ||||||||||||||||||||||||||||||||||||||||||||
Db   1629 AAATGACAGACAGATCCCTCCTATCTCCCCATAAGAGTTTGAGT 1673
```

RESULT 6

US-10-798-037-1

; Sequence 1, Application US/10798037
; Publication No. US20040229360A1
; GENERAL INFORMATION:
; APPLICANT: Robbins, Jeffrey
; APPLICANT: Children's Hospital Medical Center
; TITLE OF INVENTION: Cardiac-preferred Genetic Alteration of
; TITLE OF INVENTION: Transgenic Rabbits
; FILE REFERENCE: CHM02/GN054
; CURRENT APPLICATION NUMBER: US/10/798,037
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/454,947
; PRIOR FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5190
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus

US-10-798-037-1

Query Match   7.1%;   Score 407.6; DB 18; Length 5190;
Best Local Similarity 58.8%;   Pred. No. 1.5e-114;
Matches 1265; Conservative 0; Mismatches 699; Indels 188; Gaps 25;

```
Qy   2099 ACCTGTGCCAGAGCCGTCCAT-CTGTGTCCACATACTCTAGAATGTTCATCAGACTGCAG 2157
          ||||||||||||||||||| ||| |||| ||||  || |||||||||||||||
Db   2140 ACCTTTGCCAGAGTACGTCCGTCCCGTGCCCTGCCCTCCACTTCCAGCCAGCAGAATGCGG 2199

Qy   2158 GGCTGGCTTGGGAGGGCAGCTG------------GAAAGAGTATGTGAGAGCCAGGGG 2201
          |||| ||  ||||||||||||||              ||||||||||||||||||||||
Db   2200 AGTTGATTTGGGTGGCAGTGGGAGGACAGCGTGACAGAGAGGTGTATGGGAACGCCCAGGA 2259

Qy   2202 GAGACAGAGGGGGCCTAGGAAAGGAAGAGAGGGCGAAACCAGGCCACAAGAGGGGCAGAG 2261
          ||| ||||| |||||||||| |||||||||||||||||||||||||||||||||
Db   2260 CAGAGGAGGGGGCCCCAGGGAAG---GAAGAGGGCGAAACCAGGCCAGACCCG--GGCCCGAG 2314

Qy   2262 CCCAGAACTGAGTTAACTCCTTCCTTGTGCATCTTCCATAGAGGCAGTGGGAACTCTG 2321
          ||||||||||||| || ||||||||| |||| |||||||| || ||| ||||  || ||
Db   2315 CCCAGAGCAGGCAGGCTAACCCCTTCCTTCCTTCCCTAGGAGG--GCAGGCACTGCC 2372

Qy   2322 TGACCACCATCCCCATGAGCCCCCACTACCCACCATCCCCCAAGTTGGCCTGAGTGCATTCT 2381
          || |||||||||| ||||| ||||||||  || |||||||||||||||| ||||||||| ||
Db   2373 TGCCCACCCTCAGCCCCCTCCTACCACCTCTCCCCAAGTTTGGTCTGCTCAGGGCGCCA 2432
```

```
Qy   2382 AGGTTC-CCTGAGGACAGAGCTCGGCCTTGTCTCTTGGACCTGACCCAAGCTGACCCAA 2440
          ||||||  ||||||||||||||||||||| ||| || ||||||| ||||||||||||||
Db   2433 GGGTTCTCCCGAGGACAAAGCCCTGGTCTTTGTCACCTGACCTTGTCACCTGACCTGACCTGG 2492

Qy   2441 TGTTCTCAGTACCTTGAATGCCCTGAAGAGCTTGAGAACCAGCAGTGACATATTAGGC 2500
          ||||||||||||| ||||||||| ||||||||||| |||| ||||||||||  ||
Db   2493 TATTC-CAGCCCTTATCAGGTCCCAGGGGATGGGAACCAGGCAGTGACGACCAGGC 2551

Qy   2501 CATGGGCTAACCTGGAGCTTGCACAGAGAGCCTCAAGTGACCTCAGGACAACAGCTG 2560
          ||||||||  ||||  || |||| ||| |||||||||| |||
Db   2552 CACGGGGATAACCC-CCAGCTGTCACTGGGCCTCAAGTGACCT------AGAGCTA 2602

Qy   2561 CAGACAGGTGGCCTTTATCCCAAAGAGCAACCATTGGCATAGGTGGCTGC---AAATG 2617
          |||||| ||||||||| ||| |||| | |||||  ||||||||||||||
Db   2603 GAGACAGGTGGCCTGCATCCCCTGAGGGCAACCGTTGGCATGTGCAGGAGATG 2662

Qy   2618 GGAATGCAAGGTTGAATCAGGTCCCTTCAAGAATACTGCCGTAAGACCTAAGACCCCTG 2677
          |||||||||  |||||||||||||| ||||  ||| |||||||||||||| ||||||||
Db   2663 GGAATGGCCGGTCAACTCAGGTCCCTCCAAGGACACTGGTCTGGGGTCCAAGTCCCAA 2722

Qy   2678 GAGAGAGGGGTATGCTCTGCCCCACCCACCATAAGGGGAGTGAACTATCCTAGGGGGC 2737
          |||| |||||||||||||||||  ||||||| ||||| |||||||||||||||||||||
Db   2723 GAGAGAGGGGTGTCTCTGCCCCATCTCCCGAGCCCCG------TAATAGGCAGCCAGCCACCTAGGGGC 2774

Qy   2738 TGGCGACCTTGGGGGAGCACCA--CATTACTGAGAGTGCTGAG-CCCAGAAAAACTGACC 2794
          ||||||||| |||  ||  |||  ||||| ||||||||| ||| |||| ||||| |||||
Db   2775 TGACAACCTTGGGGGGAACCCCATCTTCACGGCGAAGTGCTCCCCACCCAGAGAAGCTGACC 2834

Qy   2795 GCCCTGTGTCCTGCCGCCACTCCACACTTCTAGAGCTATATTGAGGGTGACAGTAGATAGG 2854
          |||||| ||| ||| ||||||||| ||| || |||| |||||| |||| || || ||||
Db   2835 CCTGCCTGTCCTACCCCACCTCCACACCCTCCGGAGCTATATCGGGATAGTGCAGTAGTGGATAGG 2894

Qy   2855 GTGGGAGCTGGTAGCAGGGAGCAGTGTTCCTGGGGTGTTAGGGGGTGAGGGTGAGGGAAAGCCAGAGC 2914
          || || ||||| || |||  ||||| || | | ||  ||| ||| |  ||| || || |||||||
Db   2895 G-GTGGGGAGGGAGCTGAGCAGTTCTGGGTGTGGGAGGGTGAGGGTGAGGGAAGCCAGGGC 2953

Qy   2915 AGGGGAGTTCTGGCTTTGTCTCTGAACACAATGTCTACTTAGTTATAACAGGCATGACCT 2974
          ||||| ||| ||| || |  ||| | |||| ||| ||| ||||| |||| || || |||
Db   2954 AGGAGAGTCTGGCTTGTCTTCTGAACACAATGTCCACTTAGTCACCTAGTCACCAGCACGCGCCCT 3013

Qy   2975 GCTAAAGACCCAACATCTACGACCTCTGAAA--AGACAGCAGCCTGAGACAGGAGGGGGTT 3032
          || |||| |||| ||| ||||| ||||| ||  |||||||||| |||||| |||||| ||
Db   3014 GTTGAAGACCCGACACCTACCGCGCCTCTGAGAGGGGGGACAGCCTGCCAGGATGGGAGGAGGAGT 3073

Qy   3033 GTCTCTGAGCCTTGGGTGCTTGATGGTGTGCCACAAGGAGGGCATGAGTGTGAGTATAAGG 3092
          |||| |||||| |||||| | || || |  ||| |||| |||  |||| |||||||||||
Db   3074 ACTTGAGGTCGTCTGGGCTGCTTGAGAGCATGCATGCATGGGGAGTGTGAGC 3133

Qy   3093 CCCAGAGAGCGTTAGAGAGGGGCACTTGGGAGAGGGCACTTGGCAGAGGGGTCAGTCTGCAGAGAGCCCCCATCCAT 3152
          ||||| | | || ||||||| ||||||||| |||| || ||||  ||||||| |  || | || ||||||
Db   3134 TCAGGAGGGGATAGCGGCTTCTATCTGGAGAGGGGGCAGACTACAGAGCCCCCCTCCCCC 3193

Qy   3153 GGAATCTGAGCCGCCAGGCTGGGGCCAACTGGTGTAAATCTCTGGCGTTGCCAGGCATTCAAAGCA 3212
          ||||| || || || ||||| ||| || || |||| || || |||| |||||||| |||||||||
Db   3194 CAACCCTGCACCCCTGCAGCCCCTGTGGAACGCAGCAGAATGCGG 3253

Qy   3213 GCACCTGCATCCTCTGGCGAGCCCCCCAGCCTTCGCGGCGGAGCAACCCCCCACTTATACCCT 3272
          |||| |||| |||||||| ||||||| ||| || |||| || |||| |||||||
Db   3254 CCCCAGGCCTGCGTCTGCCGCCCATCCGGCCCCCTGCACCTTCCTCGGCAGCCTG 3313

Qy   3273 TTCTCCCCCAGCCCCAGGATTAACACTCTCTGGCCTTCCCCCCTTCCCACCTCCACCTCCATCAGGA 3332
          |||| ||| |||| |||| |||| || |||| |||| ||||||| | ||| | || |||||||
Db   3314 GAGAGGCCGAAGGGAGCGGCCGCCCACCCCCCTCCCCTGCTCCCTCAGGCCCCGGGGATTAGT 3373

Qy   3333 GTGGAGGTTGCAGAGAGGAGGTTGCAGAGGGGTAAAACCTACACGTCCAAACATCATCATGGTGCACGATAT 3392
          || ||||||| ||||| || ||| || || ||||||||||||| ||| |||| |||||||| ||
Db   3374 GCAGAAGGTGGCAGCAGGAGGGGGTAAAAGCCTTTGCGTGCA--------GATGT 3418

Qy   3393 ATGGATCAGTATGTGTAGAGGCAAGAAAGGAAATCTGCAGGCTTAACTGGGTTAATGTGT 3452
          |||| |||||| | |||| || ||| |||||| || ||||| ||||| |||| |||| |
Db   3419 GAGCATCCGCCTGCTTGGAGACGCGAGGGGATCTCTGTAAGCTGAAGTCGAAGTGTCAGAAA------ 3472
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/874,389
FILING DATE: 26-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/161,902
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-874-389-7

Query Match   5.4%;   Score 311.8; DB 9;   Length 520;
Best Local Similarity   99.4%;   Pred. No. 2.5e-85;
Matches 313; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy 4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 4336
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCAC 122

Qy 4337 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy 4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy 4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy 4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 362

Qy 4577 CGAGCTCGGTACCAG 4591
         |||||||||||||||
Db   363 CGAGCTCGGTACCCG 377

RESULT 8
US-09-921-650-7
; Sequence 7, Application US/09921650
; Publication No. US2003022315A1
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
          Gossen, Manfred
TITLE OF INVENTION: Tetracycline-Inducible Transcriptional

Qy 3453 AAAGTCTGTGTGCATGTGTGTCTGACTGAAAACGGGCATGGCTGTGCAGCTGTTCA 3512
         ||||||| ||||| ||||
Db 3473 A-------------------------TCTGCATGTGTGTCTGGTCA 3492

Qy 3513 GTTCTGTGTCGTGAGGTTACCAGACTCGAGGTTTGTGTGTAAATTGCCCAAGGCAAAGTGG 3572
         |||| |||| || ||||||||||||| ||| |||||| |||||||||||||||||
Db 3493 AGTCTGCGACGCGGGGGTCTGGACCAAGGGTCTGTGCCCAGGTGTGCCTGAGGCAGCGTG 3552

Qy 3573 GTGAATCCTTCCATGGTTTAAAGAGATTGGATGATGCCTGCATCTCAAGGACCATGGA 3632
         | || || ||||||| ||||||| |||||||||||||
Db 3553 GAGAATTCTTCCACTGGGTGATGGGTTTGCATGAGGGGCCACTGTATGGGAGCACCGGG 3612

Qy 3633 AAATAGAATGGACACTCTATATGTGTCTTAAGCTAAGTAGTAGCAAGGTCTTTGGAGGACA 3692
         | ||||||||| ||| |||||||||||| |||||||
Db 3613 GAAGTGCTGGGGCTCCAGATGCATTCCCAAGCAAAGCAAAGCCTTGAGTCCTCTGTCAGACA 3672

Qy 3693 CCTGTCTAGAGATGTGGGCAACAGAGACTACAGACAGTATCTGTACAGAGTAAGGAGAGA 3752
         |||| |||| |||| |||||||| |||||
Db 3673 CCTGCC------TGGGAATAGAAGCCCAGGCACACAGCCGTGTGCAGAGTTGGGGCAGG 3724

Qy 3753 GAGGAGGGGGGTGTAGAATTCTTCCTTACTATCAAAGGGAAACTGAGTCGTGCACCTGCAAAG 3812
         ||| |||||||||| ||||
Db 3725 ACAGAG--------------------AGGGAAACTGAGTC-TGGAGCTGGCTAG 3757

Qy 3813 TGGATGCTCTCCCTAGACATCATGACTTTGTCTCTGGGGAGCCACTGTGGAACTTCA 3872
         ||| ||||||| ||||
Db 3758 TAGACGGCTGCTCCAGCTCACAGG------------------TCCA 3785

Qy 3873 GGTCTGAGAGATTAGGAGGCTCCCCTCAGCCT-GAAGCTATGCAGATAGCCAGGGGTTGAA 3931
         |||||| ||||| ||||||||| |||||||||||||
Db 3786 GGTCTGAGAGTGTCGGGAGACACTTGCGAGCCTCGGGCTGTGTGAAGACAGACCTTGGGTAAA 3845

Qy 3932 AGGGGGAAGGGA-GAGCCTCGGGATGGAGCTTGTGTGTTGGAGGAGGCAGGGGACAGATATTA 3990
         || || || |||||| ||||
Db 3846 TGGGAGAGGGTTAGGGGGACAAGGGAGGGGAAACTGCAGCTGGGGAAACTGCAGCAAGCATTC 3905

Qy 3991 AGCCTGGAAGAGAGAAGGTGACCCTTACCCAGTTGT--TCAACTCACCCTTCAGATTAAAAA 4048
         ||| ||| |||| ||| |||| ||||||| |||
Db 3906 GTCCTATATGAAAAGGTGACCCTCACCCCCAGTGTCTCAACTCACCCTTCAGGTTAAAA 3965

Qy 4049 TAACTGAGGTAAGGGCCTGGGTAGGGAGGTGGTGT---GAGACGCTCCTGCTCTCCT 4104
         ||| ||||||||||||||||||||||||||| |||| |||||||||||
Db 3966 TAACCGAGGTAAAGGCCATGGTGTGGGGGGCGGGGAAGTGGGTGGGAGAAGTTCCTGTTCTTCCCA 4025

Qy 4105 CTGCATGCCCTGAGGCCCTTTGGGGAGGAGGAATGTGCCCAAGGACTAAAAAAAGGCCAT 4164
         ||||| |||| || ||||| ||||||| |||||||||||||||||||
Db 4026 CTATCTGCCCATCAGCGCTCTGAGGCGGCGGAATGCTCAAGGACTAAAAAAAGGCCCT 4085

Qy 4165 GGAGCCAGAGGGGCCAGAGGGCAACAGACCTTTCATGGGCAAACCTTGGGCCC 4216
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db 4086 GGAGCCCCAGGGGCGTGGGGCGCTGGGGCAGCAGACCTTTCATGGGCAAATCTGGGGGCCC 4137

RESULT 7
US-09-874-389-7
; Sequence 7, Application US/09874389
; Patent No. US2002152489A1
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
          Gossen, Manfred
TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible
                    Transcription
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

Activator Fusion Proteins

NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 28 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/921,650
 FILING DATE: 03-Aug-2001
 CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/912,650
 FILING DATE: 2001-08-03
 APPLICATION NUMBER: US 08/485,978
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/383,754
 FILING DATE: 03-FEB-1995
 APPLICATION NUMBER: US 08/275,876
 FILING DATE: 15-JULY-1994
 APPLICATION NUMBER: US 08/270,637
 FILING DATE: 01-JULY-1994
 APPLICATION NUMBER: US 08/260,452
 FILING DATE: 14-JUNE-1994
 APPLICATION NUMBER: US 08/076,327
 FILING DATE: 14-JUNE-1993
 APPLICATION NUMBER: US 08/076,726
 FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
 NAME: DeConti, Giulio A. Jr.
 REGISTRATION NUMBER: 31,503
 REFERENCE/DOCKET NUMBER: BBI-009C6CNDV
TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 520 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-921-650-7

Query Match   5.4%;   Score 311.8; DB 10; Length 520;
Best Local Similarity 99.4%; Pred. No. 2.5e-85;
Matches 313; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy 4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCAC 4336
Db   63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCAC   122

Qy 4337 TCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
Db   123 TCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG   182

Qy 4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTT 4456
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTT   242

Qy 4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG   302

Qy 4517 ATAGAGAAAAGTGAAAGTCGAGTTCCCTATCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576

Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAAGTGAAAGT 362

Qy 4577 CGAGCTCGGTACCAG 4591
Db   363 CGAGCTCGGTACCCG   377

RESULT 9
US-09-241-347-7
; Sequence 7, Application US/09241347
; Publication No. US20040003417A1
GENERAL INFORMATION:
 APPLICANT: Bujard, Hermann
          Gossen, Manfred
 TITLE OF INVENTION: Animals Transgenic for a Tetracycline-Controlled Transc
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:
  ADDRESSEE: LAHIVE & COCKFIELD
  STREET: 60 State Street, Suite 510
  CITY: Boston
  STATE: Massachusetts
  COUNTRY: USA
  ZIP: 02109-1875
 COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/241,347
  FILING DATE: 02-Feb-1999
  CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/486,814
  FILING DATE: <Unknown>
  APPLICATION NUMBER: US 08/383,754
  FILING DATE: 03-FEB-1995
  APPLICATION NUMBER: US 08/275,876
  FILING DATE: 15-JULY-1994
  APPLICATION NUMBER: US 08/270,637
  FILING DATE: 15-JULY-1994
  APPLICATION NUMBER: US 08/270,637
  FILING DATE: 01-JULY-1994
  APPLICATION NUMBER: US 08/260,452
  FILING DATE: 14-JUNE-1994
  APPLICATION NUMBER: US 08/076,327
  FILING DATE: 14-JUNE-1993
  APPLICATION NUMBER: US 08/076,726
  FILING DATE: 14-JUNE-1993
 ATTORNEY/AGENT INFORMATION:
  NAME: DeConti, Giulio A. Jr.
  REGISTRATION NUMBER: 31,503
  REFERENCE/DOCKET NUMBER: BBI-009CP4
 TELECOMMUNICATION INFORMATION:
  TELEPHONE: (617)227-7400
  TELEFAX: (617)227-5941
 INFORMATION FOR SEQ ID NO: 7:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 520 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: DNA
  SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-241-347-7

Query Match   5.4%;   Score 311.8; DB 11; Length 520;
Best Local Similarity 99.4%; Pred. No. 2.5e-85;
Matches 313; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy 4277 GGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCAC 4336
Db   63 GGATCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTCGAGTTTACCAC   122

Qy   4337 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 4396
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   123 TCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAG 182

Qy   4397 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 4456
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   183 AAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTT 242

Qy   4457 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 4516
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   243 TACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTG 302

Qy   4517 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 4576
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   303 ATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGT 362

Qy   4577 CGAGCTCGGTACCAG 4591
           |||||||||||||||
Db   363 CGAGCTCGGTACCCG 377

RESULT 10
US-10-134-643-17
; Sequence 17, Application US/10134643
; Publication No. US2003011389BA1
; GENERAL INFORMATION:
;   APPLICANT: OLSEN, JOHN C.
;   APPLICANT: MITROPHANOUS, KYRIACOS ANDREOU
;   APPLICANT: ROHLL, JONATHAN
;   APPLICANT: KINGSMAN, ALAN JOHN
;   APPLICANT: ELLARD, FIONA MARGARET
;   TITLE OF INVENTION: METHODS FOR PRODUCING HIGH TITRE VECTORS AND
;   TITLE OF INVENTION: COMPOSITIONS USED IN SUCH METHODS
;   FILE REFERENCE: 078883-0148
;   CURRENT APPLICATION NUMBER: US/10/134,643
;   CURRENT FILING DATE: 2002-08-30
;   PRIOR APPLICATION NUMBER: 60/287,048
;   PRIOR FILING DATE: 2001-04-30
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
;   LENGTH: 990
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: nucleic acid EIAV-TRE hybrid LTR
US-10-134-643-17

Query Match   5.4%; Score 311.4; DB 15; Length 990;
Best Local Similarity 99.7%; Pred. No. 4.7e-85;
Matches 312; Conservative   0; Mismatches   1; Indels   0; Gaps   0;

Qy   4279 AACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTC 4338
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   156 AACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTC 215

Qy   4339 CCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAA 4398
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   216 CCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAA 275

Qy   4399 AGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA 4458
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   276 AGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTA 335

Qy   4459 CCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGAT 4518
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   336 CCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGAT 395

Qy   4519 AGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCG 4578
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   396 AGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCG 455

Qy   4579 AGTCGGTACCAG 4591
           ||||||||||||
Db   456 AGTCGGTACCCG 468

RESULT 11
US-10-102-143-12/c
; Sequence 12, Application US/10102143
; Publication No. US2003018585A1
; GENERAL INFORMATION:
;   APPLICANT: Soldati, Dominique
;   APPLICANT: Meissner, Markus
;   TITLE OF INVENTION: TET transactivator system
;   FILE REFERENCE: 04630/016001
;   CURRENT APPLICATION NUMBER: US/10/102,143
;   CURRENT FILING DATE: 2003-01-21
;   NUMBER OF SEQ ID NOS: 31
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 4438
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1193)..(1885)
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: pTetO7Sag1-HXGPRT
US-10-102-143-12

Query Match   5.4%; Score 311; DB 16; Length 4438;
Best Local Similarity 96.9%; Pred. No. 1.4e-84;
Matches 317; Conservative   0; Mismatches 10; Indels   0; Gaps   0;

Qy   4263 CTAACGCAGCAGGAGGAACTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAA 4322
              || || |||||||||||||||||||||||||||||||||||||||||||||||
Db   979 CCTTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAA 920

Qy   4323 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCC 4382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   919 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCC 860

Qy   4383 TATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAA 4442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   859 TATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAA 800

Qy   4443 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACC 4502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   799 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGAGTTTACC 740

Qy   4503 ACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGGTACC 4562
           |||||||||||||||||||||||||||||||||||||||
Db   739 ACTCCCTATCAGTGATAGAGAAAAGTGAAAGTCGGTACC 680

Qy   4563 AGAAAAGTGAAAGTCGAGCTCGGTACC 4589
           |||||||||||||||||||||||||||
Db   679 AGAAAAGTGAAAGTCGAGCTCGGTACC 653

RESULT 12
US-10-102-143-10/c
; Sequence 10, Application US/10102143
; Publication No. US2003018585A1
; GENERAL INFORMATION:
;   APPLICANT: Soldati, Dominique
;   APPLICANT: Meissner, Markus
;   TITLE OF INVENTION: TET transactivator system
;   FILE REFERENCE: 04630/016001
;   CURRENT APPLICATION NUMBER: US/10/102,143
;   CURRENT FILING DATE: 2003-01-21
;   NUMBER OF SEQ ID NOS: 31
;   SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10

```
;   LENGTH: 4479
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1193)..(1924)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1102)
;   OTHER INFORMATION: n is disclosed as an asterisk
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pTetO7Sag1-GFP
US-10-102-143-10

Query Match             5.4%;   Score 311;  DB 16;  Length 4479;
Best Local Similarity   96.9%;          Pred. No. 1.4e-84;
Matches  317;  Conservative   0;  Mismatches  10;  Indels   0;  Gaps   0;

Qy    4263 CTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCCCTATCAGTGATAGAGAAAGTGAAA 4322
                |    ||||||||||||||||||||||||
Db     979 CTTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAA 920

Qy    4323 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCC 4382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     919 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCC 860

Qy    4383 TATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAA 4442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     859 TATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAA 800

Qy    4443 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACC 4502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     799 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACC 740

Qy    4503 ACTCCCTATCAGTGATGATGAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAG 4562
           ||||||||||||||||||||||||||||||||||||||||
Db     739 ACTCCCTATCAGTGATGATGAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAG 680

Qy    4563 AGAAAGTGAAAGTCGAGCTCGGTACC 4589
           ||||||||||||||||||
Db     679 AGAAAGTGAAAGTCGAGCTCGGTACC 653


RESULT 13
US-10-102-143-8/c
;   Sequence 8, Application US/10102143
;   Publication No. US20030185851A1
;   GENERAL INFORMATION:
;   APPLICANT: Soldati, Dominique
;   APPLICANT: Meissner, Markus
;   TITLE OF INVENTION: TET transactivator system
;   FILE REFERENCE: 04630/016001
;   CURRENT APPLICATION NUMBER: US/10/102,143
;   CURRENT FILING DATE: 2003-01-21
;   NUMBER OF SEQ ID NOS: 31
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 8
;   LENGTH: 4556
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1270)..(2001)
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: pTetO7Sag4-GFP
US-10-102-143-8

Query Match             5.4%;   Score 311;  DB 16;  Length 4556;
Best Local Similarity   96.9%;          Pred. No. 1.4e-84;
Matches  317;  Conservative   0;  Mismatches  10;  Indels   0;  Gaps   0;

Qy    4263 CTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCCCTATCAGTGATAGAGAAAGTGAAA 4322
```

```
RESULT 14
US-10-102-143-6/c
;   Sequence 6, Application US/10102143
;   Publication No. US20030185851A1
;   GENERAL INFORMATION:
;   APPLICANT: Soldati, Dominique
;   APPLICANT: Meissner, Markus
;   TITLE OF INVENTION: TET transactivator system
;   FILE REFERENCE: 04630/016001
;   CURRENT APPLICATION NUMBER: US/10/102,143
;   CURRENT FILING DATE: 2003-01-21
;   NUMBER OF SEQ ID NOS: 31
;   SOFTWARE: PatentIn Ver. 2.1
;   SEQ ID NO 6
;   LENGTH: 6346
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1193)..(3787)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1102)
;   OTHER INFORMATION: n is disclosed as an asterisk
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: pTetO7Sag1-MyoA
US-10-102-143-6

Query Match             5.4%;   Score 311;  DB 16;  Length 6346;
Best Local Similarity   96.9%;          Pred. No. 1.7e-84;
Matches  317;  Conservative   0;  Mismatches  10;  Indels   0;  Gaps   0;

Qy    4263 CTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCCCTATCAGTGATAGAGAAAGTGAAA 4322
                |    ||||||||||||||||||||||||
Db     979 CTTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAA 920

Qy    4323 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCC 4382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     919 GTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCC 860

Qy    4383 TATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAA 4442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     859 TATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAA 800

Qy    4443 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACC 4502
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     799 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACC 740

Qy    4503 ACTCCCTATCAGTGATGATGAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAG 4562
```

```
Db     739 ACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAG 680
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy    4563 AGAAAAGTGAAAGTCGAGCTCGGTACC 4589
           |||||||||||||||||||||||||||
Db     679 AGAAAAGTGAAAGTCGAGCTCGGTACC 653
```

RESULT 15
US-10-102-143-4/c
; Sequence 4, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
;   APPLICANT: Soldati, Dominique
;   APPLICANT: Meissner, Markus
;   TITLE OF INVENTION: TET transactivator system
;   FILE REFERENCE: 04630/016001
;   CURRENT APPLICATION NUMBER: US/10/102,143
;   CURRENT FILING DATE: 2003-01-21
;   NUMBER OF SEQ ID NOS: 31
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 4
;   LENGTH: 6423
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1270)..(3864)
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:
;   OTHER INFORMATION: pTetO7Sag4-MyoA
US-10-102-143-4

Query Match   5.4%;   Score 311; DB 16;   Length 6423;
Best Local Similarity 96.9%; Pred. No. 1.7e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy   4263 CTAACGGACAGGAGGGAACTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAA 4322
           | ||||||||||||||||||||||||||||||||||||||||||||
Db   979 CTTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAA 920

Qy   4323 GTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCC 4382
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   919 GTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCC 860

Qy   4383 TATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAA 4442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   859 TATCAGTGATGATAGAGAAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAA 800

Qy   4443 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACC 4502
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   799 GTGAAAGTCGAGTTTACCACTCCCTATCAGTGATGATAGAGACAAAAGTGAAAGTCGAGTTTACC 740

Qy   4503 ACTCCCTATCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAG 4562
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   739 ACTCCCTATCAGTCAGTGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAG 680

Qy   4563 AGAAAAGTGAAAGTCGAGCTCGGTACC 4589
           |||||||||||||||||||||||||||
Db   679 AGAAAAGTGAAAGTCGAGCTCGGTACC 653
```

Search completed: April 25, 2005, 15:35:15
Job time : 2159.75 secs